

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2003 年 10 月 9 日 (09.10.2003)

PCT

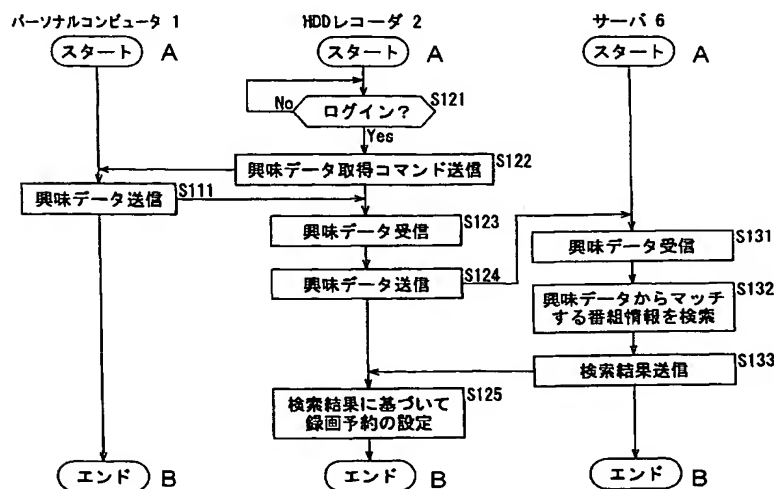
(10) 国際公開番号
WO 03/083723 A1

- (51) 国際特許分類⁷: G06F 17/30, H04N 7/173 (72) 発明者; および
(21) 国際出願番号: PCT/JP03/03795 (75) 発明者/出願人 (米国についてのみ): 山本 則行 (YAMAMOTO, Noriyuki) [JP/JP]; 〒141-0001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP). 斎藤 真里 (SAITO, Mari) [JP/JP]; 〒141-0001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP).
(22) 国際出願日: 2003 年 3 月 27 日 (27.03.2003)
(25) 国際出願の言語: 日本語
(26) 国際公開の言語: 日本語
(30) 優先権データ: 特願 2002-95414 2002 年 3 月 29 日 (29.03.2002) JP (74) 代理人: 稲本 義雄 (INAMOTO, Yoshio); 〒160-0023 東京都新宿区西新宿 7 丁目 11 番 18 号 711 ビルディング 4 階 Tokyo (JP).
(71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒141-0001 東京都品川区北品川 6 丁目 7 番 3 5 号 Tokyo (JP). (81) 指定国 (国内): CN, JP, KR, US.

[続葉有]

(54) Title: INFORMATION SEARCH SYSTEM, INFORMATION PROCESSING APPARATUS AND METHOD, AND INFORMATION SEARCH APPARATUS AND METHOD

(54) 発明の名称: 情報検索システム、情報処理装置および方法、情報検索装置および方法



1...PERSONAL COMPUTER
A...START
S111...TRANSMIT DATA ON INTEREST
B...END
2...HDD RECORDER
S121...LOG IN?
S122...TRANSMIT DATA-ON-INTEREST ACQUISITION COMMAND
S123...RECEIVE DATA-ON-INTEREST
S124...TRANSMIT DATA-ON-INTEREST
S125...SET RECORDING RESERVATION ACCORDING TO SEARCH RESULT
6...SERVER
S131...RECEIVE DATA-ON-INTEREST
S132...SEARCH PROGRAM INFORMATION MATCHED WITH DATA-ON-INTEREST
S133...TRANSMIT SEARCH RESULT

(57) Abstract: An information search system, an information processing apparatus and method, and an information search apparatus and method for recommending program information corresponding to a user's interest. A PC extracts a word corresponding to a user's interest from mail documents transmitted/received by the user and records data on interest. In S121 and S122, when log-in is performed by the user, an HDD recorder requests to acquire data on interest and according to this request, the PC transmits data on interest corresponding to the user logged in. In S123 and S124, the HDD recorder transmits the received data on interest to a server. In S131, the server searches and transmits program information matched with the data on interest received. In S125, the HDD recorder sets a program recording reservation according to the program information contained in the search result. The present invention can be applied to a program installed in a personal computer.

(57) 要約: 本発明は、ユーザの興味に対応する番組情報を推薦するようにした情報検索システム、情報処理装置および方法、情報検索装置および方法に関する。PCは、ユーザが送受信したメールの文書からユーザの興味に対応する単語を抽出し、興味データを記録する。S121、S122で、HDDレコーダは、ユーザによ

リログインされると、興味データの取得を要求し、その要求に基づいて、PCは、ログインされたユ

[続葉有]



(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

一ザに対応する興味データを送信する。S 1 2 3, S 1 2 4 で、HDDレコーダは、受信した興味データをサーバに送信し、S 1 3 1 で、サーバは、受信した興味データにマッチする番組情報を検索して送信する。S 1 2 5 で、HDDレコーダは、検索結果に含まれる番組情報に基づいて、番組の録画予約を設定する。本発明は、パーソナルコンピュータにインストールされるプログラムに適用できる。

明細書

情報検索システム、情報処理装置および方法、情報検索装置および方法

技術分野

- 5 本発明は、情報検索システム、情報処理装置および方法、情報検索装置および方法に関し、特に、例えば、電子メールなどの文書の中から、ユーザの興味があると思われる単語を取得し、その単語に関連する番組情報の推薦を行うようにした情報検索システム、情報処理装置および方法、情報検索装置および方法に関する。

10

背景技術

従来、テレビ番組やラジオ番組の推薦を行う方法として、例えば、初期興味登録方式、視聴履歴利用方式、あるいは強調フィルタリング方式などが存在する。

- 15 いずれの方法も元データとしては、電子番組ガイド (EPG : Electronic Program Guide) や Web 上などの番組情報 (番組メタデータ) が用いられており、それらの情報にマッチングすべきユーザの嗜好データの取得方法によって上記 3 方式に分類されている。

- 20 初期興味登録方式では、利用開始時、例えば、好きな番組カテゴリ (ドラマやバラエティなど)、好きなジャンル名 (ドラマや音楽など)、および好きなタレント名などをユーザに登録してもらい、それ以降、登録された情報をキーワードとして、番組メタデータとのマッチングを取り、推薦すべき番組名を取得する。

- 25 視聴履歴利用方式では、ユーザが番組を視聴する毎に、その番組に関する番組メタデータを蓄積し、所定量の視聴履歴 (番組メタデータ) が蓄積されると、その視聴履歴を解析して、推薦すべき番組名を取得する。また視聴履歴の代わりに、例えば、ハードディスクドライブに録画する装置などにおいては、ユーザによる録画予約や録画開始といった操作履歴を用いることもできる。この場合、漠然と

していた番組ではなく、ユーザの興味をより反映した情報を取得することができる。

- 5 強調フィルタリング方式では、ユーザの視聴（操作）履歴に対して、他のユーザの視聴履歴との間でマッチングを取り、当該ユーザと視聴履歴の類似する他のユーザの視聴履歴を取得する。そして、視聴履歴が類似する（好み似ている）他のユーザが視聴した番組で、当該ユーザが未だ視聴していない番組名を取得し、推薦する。

以上のような番組推薦方法を用いることによって、ユーザが興味のあると思われる番組を推薦することができる。

- 10 しかしながら、上述した番組推薦方法においては、番組メタデータからユーザの興味を抽出することになり（すなわち、テレビ番組に関する偏った興味を取得することになり）、また、番組メタデータの構造上、一般に通じ易い名称が用いられており、似たような番組が推薦されてしまうという課題があった。

- 15 すなわち、ユーザの日常の興味を反映することができず、タイムリーに有用な番組などを推薦することができなかった。

また番組推薦時、ユーザは、その番組が推薦されている理由がわからない課題があった。

発明の開示

- 20 本発明はこのような状況に鑑みてなされたものであり、ユーザが日常使用する電子メールを分析することにより、ユーザの興味に対応する単語を抽出し、抽出された単語にマッチする番組名を検索して推薦するとともに、その推薦理由を提示できるようにするものである。

- 25 本発明の第1の情報検索システムは、情報処理装置が、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段に基づいて、情報検索装置から送

信されてくる番組情報を受信する受信手段とを備え、情報検索装置が、番組情報を蓄積する蓄積手段と、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報を蓄積手段から検索する検索手段と、検索手段により検索された番組情報を情報処理装置に送信する送信手段とを備えることを特徴とする。

情報処理装置の抽出手段には、所定の情報を形態素解析して興味語に分解する形態素解析手段を含むようにすることができる。

情報処理装置には、抽出手段により抽出された興味語を用いてデータベースを作成するデータベース構築手段をさらに設けるようにすることができる。

- 10 情報処理装置には、受信手段により受信された番組情報に基づいて、番組の録画を制御する録画制御手段をさらに設けるようにすることができる。

情報処理装置には、受信手段により受信された番組情報の表示を制御する表示制御手段をさらに設けるようにすることができる。

- 15 情報検索装置の蓄積手段には、番組情報を用いて、番組に関連付けてデータベースを作成するデータベース構築手段を含むようにすることができる。

所定の情報は、文書情報、番組に関する嗜好情報、または、番組の視聴履歴に関する情報のうち、少なくとも1つを含むものとすることができる。

文書情報は、電子メールであるものとすることができる。

- 20 番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができる。

情報処理装置は、他の情報処理装置から前記所定の情報を取得することができる。

- 25 本発明の第1の情報処理装置は、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段とを備えることを特徴とする。

抽出手段には、所定の情報を形態素解析して興味語に分解する形態素解析手段を含むようにすることができる。

抽出手段により抽出された興味語を用いてデータベースを作成するデータベース構築手段をさらに設けるようにすることができる。

- 5 受信手段により受信された番組情報に基づいて、番組の録画を制御する録画制御手段をさらに設けるようにすることができる。

受信手段により受信された番組情報の表示を制御する表示制御手段をさらに設けるようにすることができる。

- 10 所定の情報は、文書情報、番組に関する嗜好情報、または、番組の視聴履歴に関する情報のうち、少なくとも1つを含むものとすることができる。

文書情報は、電子メールであるものとすることができる。

番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができる。

所定の情報は、他の情報処理装置から取得されるようにすることができる。

- 15 本発明の第1の情報処理方法は、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むこと
- 20 を特徴とする。

- 本発明の第1の記録媒体に記録されているプログラムは、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に
- 25 基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。

本発明の第1のプログラムは、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとをコンピュータに実行させる。

本発明の第1の情報検索装置は、番組に関する番組情報を蓄積する蓄積手段と、情報処理装置から送信されてきた、番組情報を取得するための興味語を受信する受信手段と、受信手段により受信された興味語に関連する番組情報を蓄積手段から検索する検索手段と、検索手段により検索された番組情報を情報処理装置に送信する送信手段とを備えることを特徴とする。

興味語は、情報処理装置において所定の情報が形態素解析されて分解された単語であるものとすることができる。

番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができる。

番組情報を分析する分析手段と、分析手段の分析結果に基づいて、番組情報のジャンルとキーワードを対応付ける辞書データを生成する辞書作成手段と、辞書作成手段により生成された辞書データに基づいて、番組情報にジャンルを付与して記憶するデータベース作成手段とをさらに備えるようにすることができる。

興味語からキーワードを抽出し、キーワードに基づいて、辞書データを検索してキーワードに対応するジャンルを取得し、ジャンルに基づいて、番組情報を検索するキーワード検索手段をさらに備えるようにすることができる。

辞書作成手段は、メタデータに含まれる単語のうち、特定のジャンルのメタデータにおいて共起率の高い単語を、そのジャンルのキーワードとして検出するキーワード検出手段を備えるようにすることができる。

辞書作成手段は、キーワードが検出される頻度をキーワードとともに記憶し、辞書データを生成するようにすることができる。

データベース作成手段は、番組情報に含まれる成分に基づいて、番組情報に含まれていない成分を補完するようにすることができる。

5 本発明の第1の情報検索方法は、番組に関する番組情報の蓄積を制御する蓄積制御ステップと、情報処理装置から送信されてきた、番組情報を取得するための興味語の受信を制御する受信制御ステップと、受信制御ステップの処理により受信が制御された興味語に関連する番組情報を検索する検索ステップと、検索ステップの処理により検索された番組情報の、情報処理装置への送信を制御する送信制御ステップとを含むことを特徴とする。

10 本発明の第2の記録媒体に記録されているプログラムは、番組に関する番組情報の蓄積を制御する蓄積制御ステップと、情報処理装置から送信されてきた、番組情報を取得するための興味語の受信を制御する受信制御ステップと、受信制御ステップの処理により受信が制御された興味語に関連する番組情報を検索する検索ステップと、検索ステップの処理により検索された番組情報の、情報処理装置への送信を制御する送信制御ステップとを含むことを特徴とする。

15 本発明の第2のプログラムは、番組に関する番組情報の蓄積を制御する蓄積制御ステップと、情報処理装置から送信されてきた、番組情報を取得するための興味語の受信を制御する受信制御ステップと、受信制御ステップの処理により受信が制御された興味語に関連する番組情報を検索する検索ステップと、検索ステップの処理により検索された番組情報の、情報処理装置への送信を制御する送信制御ステップとをコンピュータに実行させる。

20 本発明の第2の情報検索システムは、携帯端末装置が、番組を録画予約するための予約情報を生成する生成手段と、生成手段により生成された予約情報を情報処理装置に送信する第1の送信手段とを備え、情報処理装置が、携帯端末装置から送信されてきた予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段

とを備え、情報検索装置が、番組情報を蓄積する蓄積手段と、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報を蓄積手段から検索する検索手段と、検索手段により検索された番組情報を情報処理装置に送信する第2の送信手段とを備えることを特徴とする。

- 5 情報処理装置の抽出手段には、所定の情報を形態素解析して興味語に分解する形態素解析手段を含むようにすることができる。

情報処理装置には、受信手段により受信された番組情報に基づいて、番組の録画を制御する録画制御手段をさらに設けることができる。

- 10 情報検索装置の蓄積手段には、番組情報を用いて、番組に関連付けてデータベースを作成するデータベース構築手段を含むようにすることができる。

予約情報は、番組名、ジャンル名、または、出演者名のうち、少なくとも1つを含むものとすることができる。

予約情報は、電子メールであるものとすることができる。

- 15 番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができる。

本発明の第2の情報処理装置は、携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求

- 20 手段に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段とを備えることを特徴とする。

抽出手段には、所定の情報を形態素解析して興味語に分解する形態素解析手段を含むようにすることができる。

- 25 受信手段により受信された番組情報に基づいて、番組の録画を制御する録画制御手段をさらに設けるようにすることができる。

予約情報は、番組名、ジャンル名、または、出演者名のうち、少なくとも1つを含むものとすることができる。

予約情報は、電子メールであるものとすることができる。

番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができる。

- 本発明の第2の情報処理方法は、携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。
- 10 本発明の第3の記録媒体に記録されているプログラムは、携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。
- 15

- 本発明の第3のプログラムは、携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとをコンピュータに実行させる。
- 20

- 本発明の第3の情報検索システムは、情報処理装置が、電子メールを分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段による検索要求に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段とを備え、情報検索装
- 25

置が、番組情報を蓄積する蓄積手段と、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報を蓄積手段から検索する検索手段と、検索手段により検索された番組情報を情報処理装置に送信する送信手段とを備えることを特徴とする。

5 本発明の第3の情報処理装置は、電子メールを分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段による検索要求に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段とを備えることを特徴とする。

10 本発明の第3の情報処理方法は、電子メールを分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理による検索要求に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。

15 本発明の第4の記録媒体に記録されているプログラムは、電子メールを分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理による検索要求に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。

20 本発明の第4のプログラムは、電子メールの情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理による検索要求に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとをコンピュータに実行させる。

本発明の第 1 の情報検索システムにおいては、情報処理装置で、所定の情報が分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信され、情報検索装置で、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報が検索され、検索された番組情報が情報処理装置に送信される。

本発明の第 1 の情報処理装置および方法、並びにプログラムにおいては、所定の情報が分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信される。

本発明の情報検索装置および方法、並びに第 2 のプログラムにおいては、情報処理装置から送信されてきた番組情報を取得するための興味語が受信され、受信された興味語に関連する番組情報が検索され、検索された番組情報が情報処理装置に送信される。

本発明の第 2 の情報検索システムにおいては、携帯端末装置で、番組を録画予約するための予約情報が生成されて情報処理装置に送信され、情報処理装置で、携帯端末装置から送信されてきた予約情報が分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信され、情報検索装置で、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報が検索され、検索された番組情報が情報処理装置に送信される。

本発明の第 2 の情報処理装置および方法、並びに第 3 のプログラムにおいては、携帯端末装置から送信されてきた、番組を録画予約するための予約情報が分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信される。

本発明の第 3 の情報検索システムにおいては、情報処理装置で、電子メールが分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信され、情報検索装置で、情報処理装置から
5 送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報が検索され、検索された番組情報が情報処理装置に送信される。

本発明の第 3 の情報処理装置および方法、並びに第 4 のプログラムにおいては、電子メールが分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて
10 て情報検索装置から送信されてくる番組情報が受信される。

図面の簡単な説明

図 1 は、本発明を適用した番組検索システムの構成例を示す図である。

図 2 は、図 1 のパーソナルコンピュータのエージェントプログラムの機能を説明する図である。
15

図 3 は、エージェントプログラムをインストールして実行させるパーソナルコンピュータの構成例を示すブロック図である。

図 4 は、HDD レコーダの構成例を示すブロック図である。

図 5 は、図 1 のサーバのサーバプログラムの機能を説明する図である。

図 6 は、サーバプログラムをインストールして実行させるサーバの構成例を示すブロック図である。
20

図 7 は、エージェントプログラムによるデータベース作成処理を説明するフローチャートである。

図 8 は、図 7 のステップ S 1 の処理を説明するフローチャートである。

図 9 は、図 8 のステップ S 2 2 の処理を説明するフローチャートである。
25

図 10 は、話題ファイルの一例を示す図である。

図 11 は、単語ベクトルを構成する複数の単語に含まれる要素を示す図である。

図 1 2 は、図 7 のステップ S 3 の処理を説明するフローチャートである。

図 1 3 は、図 7 のステップ S 4 の処理を説明するフローチャートである。

図 1 4 は、話題単語テーブルの構成例を示す図である。

図 1 5 は、単語インデックステーブルの構成例を示す図である。

5 図 1 6 は、話題評価値テーブルの構成例を示す図である。

図 1 7 は、図 7 のステップ S 5 の処理を説明するフローチャートである。

図 1 8 は、図 7 のステップ S 9 の処理を説明するフローチャートである。

図 1 9 は、図 7 のステップ S 1 0 の処理を説明するフローチャートである。

図 2 0 は、興味データの一例を示す図である。

10 図 2 1 は、データベース更新処理を説明するフローチャートである。

図 2 2 は、データベースを更新させる条件を入力するユーザインタフェースの表示例を示す図である。

図 2 3 は、サーバプログラムによるデータベース作成処理を説明するフローチャートである。

15 図 2 4 は、番組メタデータの一例を示す図である。

図 2 5 は、番組情報の検索処理を説明するフローチャートである。

図 2 6 は、番組の推薦理由提示処理を説明するフローチャートである。

図 2 7 は、推薦理由の表示例を示す図である。

図 2 8 は、他の推薦理由の表示例を示す図である。

20 図 2 9 は、番組情報の検索処理を説明するフローチャートである。

図 3 0 は、番組の録画予約処理を説明するフローチャートである。

図 3 1 は、予約メールの一例を示す図である。

図 3 2 は、予約完了メールの一例を示す図である。

図 3 3 は、嗜好データの一例を示す図である。

25 図 3 4 は、本発明を適用した番組検索システムの構成例を示す図である。

図 3 5 は、嗜好データ取得処理を説明するフローチャートである。

図 3 6 は、番組情報検索処理を説明するフローチャートである。

図 3 7 は、嗜好データの一例を示す図である。

図 3 8 は、図 3 6 のステップ S 3 2 4 の処理を説明するフローチャートである。

図 3 9 は、辞書作成処理を説明するフローチャートである。

図 4 0 は、図 5 のデータ内容処理部の機能的構成例を示す図である。

5 図 4 1 は、図 3 9 のステップ S 3 6 2 の処理を説明するフローチャートである。

図 4 2 は、各成分に分解されたメタデータの構成例を示す図である。

図 4 3 は、ジャンル毎に集められたメタデータの構成例を示す図である。

図 4 4 は、図 3 9 のステップ S 3 6 3 の処理を説明するフローチャートである。

図 4 5 は、辞書データの構成例を示す図である。

10 図 4 6 は、データベース作成処理を説明するフローチャートである。

図 4 7 は、図 4 6 のステップ S 4 2 9 の処理を説明するフローチャートである。

発明を実施するための最良の形態

以下、図を参照して、本発明の実施の形態について説明する。

15 図 1 は、本発明を適用した番組検索システムの構成例を示す図である。この番組検索システムにおいては、インターネットなどのネットワーク 5 に、利用者端末としてのパーソナルコンピュータ 1、ハードディスクドライブ (HDD) レコーダ 2、およびデジタル携帯電話機 4 が接続されているとともに、推薦すべき番組情報 (番組メタデータ) を検索するサーバ 6 が接続されている。パーソナルコン

20 ピュータ 1 は、HDD レコーダ 2 と、例えば、Ethernet (登録商標) を介して接続されており、HDD レコーダ 2 には、テレビジョン受像機 3 が接続されている。すなわち、パーソナルコンピュータ 1、HDD レコーダ 2、およびテレビジョン受像機 3 は、1 ユーザ (または 1 家族) が所有しているものであって、それぞれが近傍に配置されている。

25 パーソナルコンピュータ 1 は、様々なアプリケーションプログラムの実行が可能な情報処理装置であって、例えば、電子メールの送受信、Web ページの閲覧、または文書作成などを行う。パーソナルコンピュータ 1 はまた、図 2 3 のフロー

チャートを参照して後述するように、電子メールの送受信により取得した文書から、ユーザの興味に対応する単語（以下、適宜、興味語と記述する）を抽出し、興味データのデータベースを作成する。

- HDD レコーダ 2 は、テレビ番組を、大容量のハードディスクドライブに録画
- 5 （記録）するとともに、ユーザからの指示に基づいて、録画したテレビ番組をテレビジョン受像機 3 に出力し、再生させる。HDD レコーダ 2 はまた、図 2 5 のフローチャートを参照して後述するように、パーソナルコンピュータ 1 から興味データを取得し、それを、ネットワーク 5 を介してサーバ 6 に送信することで、興味データにマッチする番組の推薦を受ける。
- 10 デジタル携帯電話機 4 は、図 3 0 のフローチャートを参照して後述するように、番組を録画予約するための電子メールを作成し、ネットワーク 5 を介してパーソナルコンピュータ 1 または HDD レコーダ 2 に送信し、番組の録画予約を実行させる。

- ネットワーク 5 は、公衆回線網、移動体無線通信網、ローカルエリアネットワーク、またはインターネットなどのネットワーク、デジタル衛星放送といった、
- 15 有線または無線のいずれでもよい。

- 図 1 に示した番組検索システムの例では、利用者端末として、パーソナルコンピュータ 1、HDD レコーダ 2、テレビジョン受像機 3、およびデジタル携帯電話機 4 を、それぞれ 1 つずつしか図示していないが、多数の利用者端末が接続されることは言うまでもない。
- 20

- 図 2 は、パーソナルコンピュータ 1 にインストールされて実行される、デスクトップマスケット（以下、エージェントと記述する）をデスクトップ上に表示するためのアプリケーションプログラム（以下、エージェントプログラムと記述する） 1 1、電子メールを送受信するためのアプリケーションプログラム（以下、
- 25 メーラ（mailer）と記述する） 1 2、および、文書作成または編集するためのワードプロセッサプログラム（以下、ワープロプログラムと記述する） 1 3 との関係を示す図である。

エージェントプログラム 1 1 は、処理の対象とする文書から、ユーザの興味に対応する単語を抽出し、番組を検索するための興味データ、および、処理の対象とする文書の関連情報を蓄積してデータベースを構築する蓄積部 2 1、処理の対象とする文書に対応する推薦情報をユーザに提示する提示部 2 2、および、エージェント 2 3 1（図 2 8）の表示等を制御するエージェント制御部 2 3 から構成される。

なお、蓄積部 2 1 および提示部 2 2 を、例えばインターネット上の任意のサーバに設置するようにしてもよい。

蓄積部 2 1 の文書取得部 3 1 は、メーラ 1 2 によって送受信された文書やワープロプログラム 1 3 によって編集された文書などのうち、自己が未処理の文書を取得して文書属性処理部 3 2 および文書内容処理部 3 3 に供給する。また、文書取得部 3 1 は、ユーザによって HDD レコーダ 2 に初期登録された嗜好情報（好みのジャンル名や出演者名など）、または視聴履歴を取得し、それらを文書内容処理部 3 3 に供給する。

15 なお、以下においては、主に、メーラ 1 2 によって送受信された電子メールの文書を処理の対象とする場合の例について説明する。

文書属性処理部 3 2 は、文書取得部 3 1 から供給される文書の属性情報を抽出し、その属性情報に基づいて文書をグループ化し、文書内容処理部 3 3 および文書特徴データベース作成部 3 4 に供給する。電子メールの場合、属性情報としては、文書のヘッダに記述されている情報（対象となっている電子メールを特定するメッセージ ID、参照している電子メールのメッセージ ID (References, In-Reply-To)、宛先 (To, Cc, Bcc)、あるいは送信元 (From)、日付 (Date))、表題 (subject) などが抽出される。そして、抽出された属性情報に基づいて、1 以上の文書がグループ化される。以下、属性情報に基づいてグループ化された文書群（電子メールグループ）を「話題」と記述する。

25 また、一般にここで言う話題とは、電子メールに限らず、ワープロ、エディタやスケジューラなど、その他のツールやアプリケーションソフトウェアなどから

作成されるようなあらゆる文書に関して、ある関係で関連付けられた一連の文書群を指す。

文書内容処理部 33 は、文書属性処理部 32 でグループ化された文書群（話題）の本文を抽出し、形態素解析を施して、単語（特徴語）に分類する。また、

- 5 文書内容処理部 33 は、文書取得部 31 から供給された嗜好情報または視聴履歴に形態素解析を施して、単語（興味語）に分類する。

単語は、品詞（名詞、形容詞、動詞、副詞、接続詞、感動詞、助詞、および助動詞）別に分類される。ただし、広範囲に亘って分布している単語、すなわち、例えば、大多数の文書に含まれていると考えられる単語「こんにちは」、「よろしく」、あるいは「お願いします」等の名詞以外の品詞は関連情報を検索するためのキーワード（以下、検索語とも記述する）には成り得ないので、不要語であるとしてキーワードとする対象から削除される。

- 15 また、文書内容処理部 33 は、不要語が削除された各単語の出現頻度および複数の文書に亘る分布状況を求め、グループ化された文書群（話題）毎に、各単語の重み（文書の主旨に関係する程度を示す値、以下、評価値と記述する）を演算する。

さらに、文書内容処理部 33 は、各話題に対し、単語の評価値を要素とする特徴ベクトルを決定する。例えば、各話題に含まれる単語（特徴語）の総数が n 個である場合、各話題の特徴ベクトルは、 n 次元空間のベクトルとして次式（1）

- 20 のように表現される。

特徴ベクトル

$$= (\text{単語 1 の評価値 } w_1, \text{ 単語 2 の評価値 } w_2, \dots, \text{ 単語 } n \text{ の評価値 } w_n) \dots (1)$$

- 25 評価値の演算には、例えば文献 (Salton, G.: Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer, Addison-Wesley, 1989) に開示されている tf-idf 法を用いる。tf-idf 法によれば、話題 A に対応する n 次元の特徴ベクトルのうち、話題 A に含ま

れる単語に対応する要素に対しては、評価値として0以外の値が算出され、話題Aに含まれない単語（頻度が0である単語）に対応する要素に対しては、評価値として0が算出される。

5 なお、評価値は、例えば、電子メールの送受信の頻度や回数、電子メールに含まれる単語の品詞の種類（特定の地域や名称を示す固有名詞など）、送受信する相手に応じて修正される。

10 また、本実施の形態においては、話題毎に特徴ベクトルを算出するものとして説明するが、これに限らず、文書毎、または、その他の単位毎（例えば、所定期間（1週間）に蓄積された文書群毎）に特徴ベクトルを算出するようにすることも勿論可能である。

15 文書特徴データベース作成部34は、文書属性処理部32によってグループ化された文書群（話題）毎の各文書の属性情報と、文書内容処理部33で算出された話題毎の特徴ベクトル（すなわち、話題内に含まれる単語の評価値）を時系列順にデータベース化するとともに、文書内容処理部33で算出された、嗜好情報
20 または視聴履歴から興味データ（後述）を作成し、それらをデータベース化し、ハードディスクドライブなどよりなる記憶部59に記録する。

25 また、文書特徴データベース作成部34は、単語の評価値などを参照することにより、所定の条件を満たす単語を選択し、関連情報や番組情報を検索するための検索用キーワード（検索語または興味語）として記録する。さらに、文書特徴データベース作成部34は、検索語を関連情報検索部35に供給し、それに対応して関連情報検索部35から供給される関連情報を、検索語に対応付けて記録する。

30 関連情報検索部35は、文書特徴データベース作成部34から供給される検索語に対する関連情報を検索し、検索結果のインデックスを文書特徴データベース作成部34に供給する。検索語に対する関連情報を検索する方法としては、例えばインターネット上の検索エンジンを用いる方法がある。検索エンジンを用いる方法を適用した場合、検索結果として得られるWebページのURL（Uniform

Resource Locator) と Web ページのタイトルが、関連情報として文書特徴データベース作成部 3 4 に供給される。

提示部 2 2 のイベント管理部 4 1 は、メーラ 1 2 がアクティブとされること、メーラ 1 2 が電子メールの送受信を完了したこと、および、入力中の文書のテキストデータ量が所定の閾値を超えたことを検知して、データベース問い合わせ部 4 2 に通知する。以下、メーラ 1 2 が電子メールの送受信を完了したこと、または、入力中の文書のテキストデータ量が所定の閾値を超えたことを、イベント発生と記述する。

また、イベント管理部 4 1 は、内蔵するタイマ 4 1 A を参照することによって時間の経過を監視し、適宜、所定のタイミングから所定の時間が経過した場合、その旨をデータベース問い合わせ部 4 2 に通知する。

データベース問い合わせ部 4 2 は、イベント管理部 4 1 からのイベント発生の通知に対応して、イベント発生に対応する文書（例えば、受信した電子メール）を取得し、文書内容処理部 3 3 の処理と同様に、その文書に形態素解析を施して単語を抽出し、不要語を除外して各単語の評価値を演算する。これにより、イベント発生に対応する文書の特徴ベクトルが算出される。

また、データベース問い合わせ部 4 2 は、文書特徴データベース作成部 3 4 によって作成されたデータベースを検索し、算出したイベント発生に対応する文書の特徴ベクトルと、データベースに記録されている話題毎の特徴ベクトルとの内積を、両者の類似度として算出する。さらに、データベース問い合わせ部 4 2 は、イベント発生に対応する文書に対する類似度が最も高い話題を判定し、その話題に含まれる単語のうち、評価値が所定の条件（詳細については後述する）を満たすものを選択し、選択した単語（重要語）に対応する関連情報（推薦情報）を、イベント管理部 4 1 を介し、または直接的に、関連情報提示部 4 3 に供給する。

さらに、データベース問い合わせ部 4 2 は、HDD レコーダ 2 にログインされたユーザに対応する興味データをデータベースから読み出して、HDD レコーダ 2 に送信したり、あるいは、イベント発生に対応して、嗜好情報または視聴履歴から

作成された興味データをデータベースから読み出して、ネットワーク 5 を介してサーバ 6 に送信し、その興味データにマッチする番組情報の検索を要求する。

関連情報提示部 4 3 は、イベント管理部 4 1 を介し、または直接的に、データベース問い合わせ部 4 2 から供給される関連情報（推薦情報）を表示部 5 8（デスクトップ）上に表示させる。すなわち、イベント管理部 4 1 がイベント発生を検知する毎、提示部 2 2 による関連情報の提示が更新される。

なお、蓄積部 2 1 によるデータベースの更新は、所定のタイミングにおいて実行される。データベースの更新処理は、図 2 1 のフローチャートを参照して後述する。また、蓄積部 2 1 によるデータベースの更新時には、記憶部 5 9 に記録した特徴ベクトルが、例えば、電子メールの送受信の頻度や回数、電子メールに含まれる単語の品詞の種類（特定の地域や名称を示す固有名詞など）に応じて修正される。

図示は省略するが、HDD レコーダ 2 にインストールされて実行されるエージェントプログラムも、上述した図 2 のエージェントプログラム 1 1 とほぼ同様の機能を有する。なお、HDD レコーダ 2 が、パーソナルコンピュータ 1 の蓄積部 2 1 を使用（共有）することで、その部分を省略するようにしてもよい。

図 3 は、エージェントプログラム 1 1 乃至ワープロプログラム 1 3 がインストールされて実行されるパーソナルコンピュータ 1 の構成例を示している。なお、当然ながら、本発明はパーソナルコンピュータの他、ホームサーバシステム、ゲーム機器、カーナビゲーションシステム、PDA (Personal Digital Assistant) 等の情報電子機器において利用できる。

このパーソナルコンピュータ 1 は、CPU (Central Processing Unit) 5 1 を内蔵している。CPU 5 1 には、バス 5 4 を介して入出力インタフェース 5 5 が接続されている。入出力インタフェース 5 5 には、キーボード、マウスなどの入力デバイスよりなる入力部 5 6、処理結果としての例えば音声信号を出力する出力部 5 7、処理結果としての画像を表示するディスプレイなどよりなる表示部 5 8、プログラムや構築されたデータベースなどを格納するハードディスクドライブな

どよりなる記憶部 5 9、インターネットに代表されるネットワークを介してデータを通信する LAN (Local Area Network) カードなどよりなる通信部 6 0、および、磁気ディスク 6 2、光ディスク 6 3、光磁気ディスク 6 4、または半導体メモリ 6 5 などの記録媒体に対してデータを読み書きするドライブ 6 1 が接続されている。バス 5 4 には、ROM (Read Only Memory) 5 2 および RAM (Random Access Memory) 5 3 が接続されている。

本発明のエージェントプログラム 1 1 は、磁気ディスク 6 2 乃至半導体メモリ 6 5 に格納された状態でパーソナルコンピュータに供給され、ドライブ 6 1 によって読み出されて、または通信部 6 0 がネットワークを介して取得して、記憶部 5 9 に内蔵されるハードディスクドライブにインストールされている。記憶部 5 9 にインストールされているエージェントプログラム 1 1 は、入力部 5 6 に入力されるユーザからのコマンドに対応する CPU 5 1 の指令によって、記憶部 5 9 から RAM 5 3 にロードされて実行される。なお、パーソナルコンピュータ 1 の起動時において自動的にエージェントプログラム 1 1 が実行されるように設定することも可能である。

また記憶部 5 9 に内蔵されるハードディスクドライブには、エージェントプログラム 1 1 の他、メーラ 1 2、およびワープロプログラム 1 3、WWW (World Wide Web) ブラウザなどのアプリケーションプログラムもインストールされており、エージェントプログラム 1 1 と同様に、入力部 5 6 に入力されるユーザからの起動コマンドに対応する CPU 5 1 の指令によって、記憶部 5 9 から RAM 5 3 にロードされて実行される。

図 4 は、HDD レコーダ 2 の構成例を示すブロック図である。この HDD レコーダ 2 は、膨大な記憶容量を備えるハードディスクドライブ (HDD) 7 8 に数多くの映像を録画することできるとともに、ユーザの意図を的確に把握して録画映像の記録管理 (視聴履歴や操作履歴など) を反映することができる。なお、HDD レコーダ 2 は、AV 機器として実装することができ、例えば、セットトップボックス (STB) のようなテレビジョン受像機と一体で構成することができる。

CPU 7 1 は、HDD レコーダ 2 全体の動作を制御するメインコントローラであり、入力部 7 6 から供給された入力信号に基づいて、チューナ 7 9、復調部 8 0、デコーダ 8 1、および HDD 7 8などを制御することにより、番組の録画または再生を実行させる。

- 5 RAM 7 3 は、CPU 7 1 の実行プログラムをロードしたり、実行プログラムの作業データを書き込むために使用される、書き込み可能な揮発性メモリである。

ROM 7 2 は、HDD レコーダ 2 の電源オン時に実行する自己診断・初期化プログラムや、ハードウェア操作の制御コードなどを格納する読み出し専用メモリである。

- 10 入力部 7 6 は、例えば、リモートコマンド、ボタン、スイッチ、あるいはキーボードなどにより構成され、操作に対応する入力信号を、入出力インターフェース 7 5 およびバス 7 4 を介して CPU 7 1 に出力する。

- 通信部 7 7 は、ネットワーク 5 を介して、サーバ 6 と通信することにより、推薦された番組メタデータを受信するとともに、パーソナルコンピュータ 1 と通信
15 することにより、所定データ（例えば、興味データ）を送受信することができる。通信部 7 7 に入力されたデータは、適宜、入出力インターフェース 7 5 を介して HDD 7 8 に記録される。

- HDD 7 8 は、プログラムやデータなどを所定フォーマットのファイル形式で蓄積することができる、ランダムアクセスが可能な記憶装置であり、膨大な記憶容
20 量を備えている。HDD 7 8 は、入出力インターフェース 7 5 を介してバス 7 4 に接続されており、デコーダ 8 1、または通信部 7 7 から、放送番組、および EPG データなどのデータ放送用のデータの入力を受け、これらの情報を記録するとともに、必要に応じて、記録されたデータを出力する。

- 図示せぬアンテナで受信された放送波は、チューナ 7 9 に供給される。放送波
25 は、所定のフォーマットに基づいており、例えば、EPG データを含んでいても良い。放送波は、衛星放送波、地上波、有線、または無線の区別を特に問わない。

チューナ 7 9 は、CPU 7 1 の制御に基づいて、所定チャンネルの放送波のチューニングすなわち選局を行い、受信データを復調部 8 0 に出力する。なお、送信されてくる放送波がアナログであるか、あるいは、デジタルであるかに応じて、チューナ 7 9 の構成を適宜変更または拡張することができる。復調部 8 0 は、デジタル変調されている受信データを復調し、デコーダ 8 1 に出力する。

例えば、デジタル衛星放送の場合、チューナ 7 9 により受信され、復調部 8 0 により復調されたデジタルデータは、MPEG2 (Moving Picture Experts Group 2) 方式で圧縮された AV データ、および、データ放送用のデータが多重化されているトランスポートストリームである。前者の AV データは、放送番組本体を構成する映像データおよび音声データであり、後者のデータ放送用のデータは、この放送番組本体に付随するデータ (例えば、EPG データ) を含むものである。

デコーダ 8 1 は、復調部 8 0 より供給されたトランスポートストリームを、MPEG 方式で圧縮された AV データとデータ放送用のデータ (例えば、EPG データ) に分離する。分離されたデータ放送用のデータは、バス 7 4 および入出力インターフェース 7 5 を介して HDD 7 8 に供給され、記録される。

受信した番組をそのまま出力することが指示されている場合、デコーダ 8 1 は、AV データを、圧縮映像データと圧縮音声データにさらに分離する。分離された音声データは、デコードされた後、ミキサ 8 3 を介してテレビジョン受像機 3 のスピーカに出力される。また、分離された映像データは、伸張された後、コンポザー 8 2 を介してテレビジョン受像機 3 のモニタに出力される。

受信した番組を HDD 7 8 に録画することが指示されている場合、デコーダ 8 1 は、分離する前の AV データを、バス 7 4 および入出力インターフェース 7 5 を介して HDD 7 8 に出力する。また、HDD 7 8 に録画されている番組の再生が指示されている場合、デコーダ 8 1 は、入出力インターフェース 7 5 およびバス 7 4 を介して、HDD 7 8 から AV データの入力を受け、圧縮映像データと圧縮音声データに分離し、それぞれコンポザー 8 2 もしくはミキサ 8 3 に出力する。

コンポーザ 8 2 は、必要に応じて、デコーダ 8 1 より入力された映像データと、GUI (Graphical User Interface) 画面とを合成して、テレビジョン受像機 3 のモニタに出力する。

図 5 は、サーバ 6 にインストールされて実行されるサーバプログラム 1 0 1 の機能の説明する図である。

サーバプログラム 1 0 1 は、処理の対象とする、EPG データなどの番組メタデータを分析し、推薦番組用データベースを構築する蓄積部 1 1 1、および、蓄積部 1 1 1 に蓄積されている推薦番組用データベースの中から、ユーザの興味データにマッチする番組情報を検索する検索部 1 1 2 から構成される。

10 蓄積部 1 1 1 の番組メタデータ取得部 1 2 1 は、EPG データなどの番組メタデータのうち、自己が未処理の番組メタデータを、図示せぬ EPG データ提供装置から取得してデータ内容処理部 1 2 2 に供給する。

データ内容処理部 1 2 2 は、番組メタデータ取得部 1 2 1 から供給される番組メタデータに形態素解析を施し、番組情報（番組名、ジャンル名、放送局、時間帯、出演者、およびキーワードなど）を抽出する。抽出された番組情報は、データベース作成部 1 2 3 に供給される。

データベース作成部 1 2 3 は、データ内容処理部 1 2 2 によって抽出された番組情報を番組毎にデータベース化して、ハードディスクドライブなどよりなる記憶部 1 4 7（図 6）に記録する。

20 検索部 1 1 2 のイベント管理部 1 3 1 は、ネットワーク 5 を介して利用者端末装置（パーソナルコンピュータ 1 または HDD レコーダ 2）から興味データの入力を検知して、データベース問い合わせ部 1 3 2 に通知する。以下、興味データの入力を検知したことを、検索要求と記述する。また、イベント管理部 1 3 1 は、内蔵するタイマ 1 3 1 A を参照することによって時間の経過を監視し、適宜、所定のタイミングから所定の時間が経過した場合、その旨をデータベース問い合わせ部 1 3 2 に通知する。

データベース問い合わせ部 1 3 2 は、イベント管理部 1 3 1 からの検索要求の通知に基づいて、検索要求に対応する興味データを取得する。データベース問い合わせ部 1 3 2 は、検索エンジンを用いて、データベース作成部 1 2 3 によって作成された推薦番組用データベースを検索し、取得した興味データにマッチする

5 番組情報を推薦用番組として選択する。

また、データベース問い合わせ部 1 3 2 は、利用者端末装置で初期登録された嗜好情報（好みのジャンル名や出演者名など）、または視聴履歴などを取得し、データ内容処理部 1 2 2 の処理と同様に、嗜好情報や視聴履歴に形態素解析を施して、興味データを抽出し、データベース作成部 1 2 3 の推薦番組用データベース

10 スを検索し、抽出した興味データにマッチする番組情報を推薦用番組として選択する。

以上のようにして選択された推薦用番組は、イベント管理部 1 3 1 を介し、または直接的に、番組情報出力部 1 3 3 に供給される。

番組情報出力部 1 3 3 は、イベント管理部 1 3 1 を介し、または直接的に、データベース問い合わせ部 1 3 2 から供給される推薦用番組（番組情報）を、ネットワーク 5 を介して利用者端末装置（パーソナルコンピュータ 1 または HDD レコーダ 2）に検索結果として出力する。

15

なお、蓄積部 1 1 1 による推薦番組用データベースの更新は、EPG データが更新される毎、あるいは、所定のタイミングにおいて実行される。

図 6 は、サーバプログラム 1 0 1 がインストールされて実行されるサーバ 6 の構成例を示している。同図に示される CPU 1 4 1 乃至半導体メモリ 1 5 3 は、図 3 に示したパーソナルコンピュータ 1 の CPU 5 1 乃至入力部 5 6 並びに記憶部 5 9 乃至半導体メモリ 6 5 と基本的に同様の構成を有するものであるため、その説明は省略する。

20

次に、パーソナルコンピュータ 1 のエージェントプログラム 1 1 によるデータベース作成処理について、図 7 のフローチャートを参照して説明する。このデータベース作成処理は、エージェントプログラム 1 1 が実行する処理のうちの 1 つ

25

であり、エージェントプログラム 11 が起動された状態において、データベースが未だ作成されていないときに開始される。

5 ステップ S 1 において、文書取得部 31 は、データベース作成の素として分析する文書（例えば、エージェントプログラム 11 が実行される以前に送受信された電子メール、以下、分析対象電子メールと記述する）を、記憶部 59 に内蔵されるハードディスクドライブから選択的に取得して文書属性処理部 32 および文書内容処理部 33 に供給する。

 ステップ S 1 の処理、すなわち、分析対象電子メール選択処理の詳細について、図 8 を参照して説明する。

10 ステップ S 2 1 において、文書取得部 31 は、ユーザが送信した電子メールが保存されている送信フォルダを参照し、直近の所定期間（例えば、最近の一週間）に送信した電子メールの数が所定数（例えば、100 通）以上存在するか否かを判定する。直近の所定期間に送信した電子メールの数が所定数以上存在すると判定された場合、処理はステップ S 2 2 に進む。ステップ S 2 2 において、文
15 書取得部 31 は、日時条件およびアドレス属性条件を設定する。

 ステップ S 2 2 の処理、すなわち、日時条件およびアドレス属性条件を設定処理の詳細について、図 9 を参照して説明する。ステップ S 3 1 において、文書取得部 31 は、送信フォルダに存在する電子メールの数が所定数（例えば、1000 通）以上であるか否かを判定する。

20 ステップ S 3 1 において、送信フォルダに存在する電子メールの数が所定数以上であると判定された場合、処理はステップ S 3 2 に進む。ステップ S 3 2 において、文書取得部 31 は、分析対象電子メールを選択するための日時条件を「1 年以前を除去」に設定する。ステップ S 3 3 において、文書取得部 31 は、分析対象電子メールを選択するためのアドレス属性条件を「” To ” 以外を除去」に
25 設定する。また、文書取得部 31 は、アドレス条件（アドレスリスト）を抽出する対象を送信フォルダに設定する。

反対に、ステップ S 3 1 において、送信フォルダに存在する電子メールの数が所定数よりも少ないと判定された場合、処理はステップ S 3 4 に進む。ステップ S 3 4 において、文書取得部 3 1 は、日時条件を「3 年以前を除去」に設定する。ステップ S 3 5 において、文書取得部 3 1 は、アドレス属性条件を「" To, C
5 c" 以外を除去」に設定する。また、文書取得部 3 1 は、アドレス条件を抽出する対象を送信フォルダおよび受信フォルダに設定する。

以上のような日時条件およびアドレス属性条件設定処理により、送信した電子メールの数に対応して、分析対象電子メールの日時条件とアドレス属性条件が設定された後、処理は、図 8 のステップ S 2 3 にリターンする。

- 10 なお、日時条件およびアドレス属性条件設定処理は、上述した 2 種類の選択だけでなく、例えば、送信フォルダのメール数に応じていくつかの区間を設け、それに応じて、日時条件を任意の年数で細かく区切ったり、受信簿に対するアドレス属性条件にさらに from, reply to 等を加えた選択肢を増やすなどしてもよい。

- 15 ステップ S 2 3 において、文書取得部 3 1 は、送信フォルダ（または受信フォルダ）に存在する電子メールを、ステップ S 2 2 で設定した日時条件およびアドレス属性条件に基づいてフィルタリングすることにより、電子メールの数を絞り込む。ステップ S 2 4 において、文書取得部 3 1 は、ステップ S 2 3 でフィルタリングされた各電子メールの宛先（または送信元）をリスト化するとともに、各宛先の出現回数をカウントし、出現回数が多い上位 n 個のアドレスを判定して、
20 アドレス条件を「上位 n 個のアドレスから送受信された電子メールを抽出」に設定する。

- 25 ステップ S 2 5 において、文書取得部 3 1 は、全ての電子メール、すなわち、送信フォルダ、受信フォルダ、およびその他のフォルダの存在する電子メールのうち、ステップ S 2 2 で設定した日時条件およびステップ S 2 4 で設定したアドレス条件に基づいてフィルタリングすることにより、分析対象電子メールを選択する。

なお、ステップS 2 1において、ユーザが送信した電子メールが保存されている送信フォルダを参照し、直近の所定期間に送信した電子メールの数が所定数よりも少ないと判定された場合、処理はステップS 2 6に進む。ステップS 2 6において、文書取得部3 1は、ユーザが送信した電子メールが保存されている受信5 フォルダを参照し、直近の所定期間（例えば、最近の一週間）に受信した電子メールの数が所定数（例えば、1 0 0通）以上存在するか否かを判定する。直近の所定期間に受信した電子メールの数が所定数以上存在すると判定された場合、処理はステップS 2 2に進み、それ以降の処理が繰り返される。

10 反対に、ステップS 2 6において、直近の所定期間に受信した電子メールの数が所定数よりも少ないと判定された場合、この段階でデータベース作成処理は終了される。

以上のように分析対象電子メールが選択された後、処理は、図7のステップS 2にリターンする。

15 ステップS 2において、文書属性処理部3 2は、ステップS 1の処理で文書取得部3 1から供給された分析対象電子メールから属性情報（メッセージID等のヘッダ情報）を抽出し、その属性情報に基づき、分析対象電子メールを話題毎に分類して（すなわち、話題毎にグループ化して）、話題毎に話題ファイルを生成して文書内容処理部3 3および文書特徴データベース作成部3 4に供給する。

20 図10は、ステップS 2において作成される話題ファイル1 6 1の一例を示している。話題ファイル1 6 1は、各話題ファイルを識別するためのトピックスID 1 6 2、当該話題に属する最古の電子メールの通信時間を示す日時情報1 6 3、当該最古の電子メールの題名などを示すサブジェクト情報1 6 4、当該話題に属する電子メールの送信元または宛先の電子メールアドレスからなるメンバー情報1 6 5、当該話題に属する各電子メールを特定するメールメッセージID 1 6 6、
25 当該話題に属する電子メールの本文に含まれる単語から構成される単語ベクトル1 6 7、当該話題に属する電子メールの本文を連結した連結本文1 6 8、および

いずれかの話題に含まれる全ての単語の評価値から成る特徴ベクトル 1 6 9 から構成される。

トピックス I D 1 6 2 として、例えば、当該話題に属する最古の電子メールの通信時間を用いるようにしてもよい。

- 5 なお、連結本文 1 6 8 は、当該話題に属する電子メールのうち、送信フォルダに存在する電子メールの本文を連結した後、所定の文字列（例えば” soshin-shuryo”）を挿入して、受信フォルダやその他のフォルダに存在する電子メールの本文を連結するようにする。

- 10 図 1 1 は、単語ベクトル 1 6 7 を構成する複数の単語 1 7 0 に含まれる要素を示している。すなわち、単語 1 7 0 には、当該単語自身の文字列 1 7 1、当該単語の品詞（名詞の種類） 1 7 2、当該話題における当該単語の頻度 1 7 3、および当該話題における当該単語の評価値 1 7 4 を記録するための構成を有している。なお、単語 1 7 0 の各要素の中身は、ステップ S 2 の処理段階では生成されず、以降の処理において生成される。

- 15 また、特徴ベクトル 1 6 9 も、ステップ S 2 の処理段階では生成されず、以降の処理において生成される。

図 7 に戻る。ステップ S 3 において、文書属性処理部 3 2 は、ステップ S 2 で生成した話題を選抜する。ステップ S 3 の処理、すなわち第 1 次話題選抜処理について、図 1 2 のフローチャートを参照して説明する。

- 20 ステップ S 4 1 において、文書属性処理部 3 2 は、ステップ S 2 で生成した話題の数が所定数以上存在するか否かを判定する。生成した話題の数が所定数以上存在すると判定された場合、処理はステップ S 4 2 に進む。ステップ S 4 2 において、文書属性処理部 3 2 は、生成した話題を選抜するための構成メール数条件を「a（例えば 4）通以下を削除」に設定する。

- 25 反対に、ステップ S 4 1 において、生成した話題の数が所定数よりも少ないと判定された場合、処理はステップ S 4 3 に進む。ステップ S 4 3 において、文書

属性処理部 3 2 は、生成した話題を選抜するための構成メール数条件を「b (a よりも小さい数、例えば 2) 通以下を削除」に設定する。

- ステップ S 4 4 において、文書属性処理部 3 2 は、上段の処理で設定した構成メール数条件に基づき、ステップ S 2 で生成した話題をフィルタリングする。すなわち、例えば、上段の処理で構成メール数条件を「a 通 (例えば 4 通) 以下を削除」に設定した場合、4 通以下の電子メールから構成される話題を削除し、5 通以上の電子メールから構成される話題だけを選抜する。

さらに、直近の所定期間 (例えば、最近の一週間) に通信した電子メールを含まない話題を削除するようにしてもよい。

- 10 このようにして第 1 次話題選抜処理を実行した後、処理は、図 7 のステップ S 4 にリターンする。

なお、第 1 次話題選抜処理における構成メール数条件の設定は、上述した 2 種類の選択だけでなく、例えば、話題の数に応じていくつかの区間を設けて、その区間ごとに構成メール数条件を決定するようにしてもよい。

- 15 ステップ S 4 において、文書内容処理部 3 3 は、選抜された各話題に対応する話題ファイル 1 6 1 の連結本文 1 6 8 に形態素解析を実行する。ステップ S 4 における形態素解析処理の詳細について、図 1 3 のフローチャートを参照して説明する。

- 20 ステップ S 5 1 において、文書内容処理部 3 3 は、選抜された各話題のうち、形態素解析を施していないものが存在するか否かを判定する。形態素解析を施していないものが存在すると判定された場合、処理はステップ S 5 2 に進む。ステップ S 5 2 において、文書内容処理部 3 3 は、形態素解析を施していない話題を 1 つ選択し、対応する話題ファイル 1 6 1 の連結本文 1 6 8 を読み出して形態素解析を施し、連結本文 1 6 8 に含まれる単語を抽出する。

- 25 このように、話題ファイル 1 6 1 の連結本文 1 6 8 に対して形態素解析を施す処理は、話題ファイル 1 6 1 を構成する電子メールの各本文に対して形態素解析

を施す処理に比較して、処理する文章は長くなるが処理回数が1回で済むので、処理に要するリソースの浪費を抑止することができる。

ステップS53において、文書内容処理部33は、ステップS52で抽出した単語のうち、品詞が名詞（一般名詞、サ変接続名詞、地名、人名、興味がある用
5 語を含む）であるものを抽出する。ステップS54において、文書内容処理部33は、抽出した名詞である単語を並べ、当該話題に対応する単語ベクトル167を生成する。

ステップS55において、文書内容処理部33は、話題単語テーブル181
（図14）にステップS54で生成した単語ベクトル167に対応する記録を追
10 加するとともに、ステップS54で生成した単語ベクトル167を構成する単語の記録を、話題評価値テーブル193を含む単語インデックステーブル191
（図15）に追加する。なお、話題単語テーブル181、単語インデックステーブル191、および話題評価値テーブル193は、いずれもハッシュテーブル（Hash table）である。

15 図14は、話題単語テーブル181の構成例を示している。話題単語テーブル181は、各話題に対するトピックスID162と、それに対応する単語ベクトル167が記録されており、トピックスID162を入力として、対応する単語ベクトル167を出力する。

図15は、単語インデックステーブル191の構成例を示している。単語イン
20 デックステーブル191は、各単語ベクトル167を構成する単語名192と、それに対応する話題評価値テーブル193の組が複数記録されており、単語名192を入力として、話題評価値テーブル193を出力する。

図16は、話題評価値テーブル193の構成例を示している。話題評価値テ
ブル193は、単語名192に対応する単語が含まれる話題のトピックスID2
25 01と、当該話題における当該単語の評価値202が記録されており、トピックスID201を入力として、当該話題における当該単語の評価値202を出力する。

このような構成の話題単語テーブル 181 乃至話題評価値テーブル 193 を生成することにより、トピックス ID 162 および単語名 192 のどちらか一方を入力としても、対応する他方を容易に検索することが可能となる。

この後、処理はステップ S 51 に戻り、以降の処理が繰り返される。その後、
5 ステップ S 51 において、選抜された各話題のうち、形態素解析を施していないものが存在しないと判定された場合、形態素解析処理は終了され、処理は、図 7 のステップ S 5 にリターンする。

ステップ S 5 において、文書内容処理部 33 は、以降における処理を軽減するために、これまでの処理で抽出された単語、すなわち、各話題にそれぞれ対応する
10 単語ベクトルに含まれる単語のうち、話題の内容に関連が薄いと考えられる単語、あいさつなどの日常的な単語等（以下、不要語と記述する）を除去する。

ステップ S 5 における不要語削除処理について、図 17 のフローチャートを参照して説明する。ステップ S 61 において、文書内容処理部 33 は、単語ベクトルが小さい話題、すなわち、対応する単語ベクトルを構成する単語の数が所定数
15 （例えば、5 個）以下である話題を除去する。

ステップ S 62 において、文書内容処理部 33 は、ステップ S 4 の処理で生成した単語インデックステーブル 191 に記録されている単語のうち、以降の処理の対象としていない単語が存在するか否かを判定する。処理対象としていない単語が存在すると判定された場合、処理はステップ S 63 に進む。ステップ S 63
20 において、文書内容処理部 33 は、単語インデックステーブル 191 に記録されている、処理対象としていない単語のうちの 1 つを処理対象の単語に選択する。

ステップ S 64 において、文書内容処理部 33 は、処理対象の単語を入力として、単語インデックステーブル 191 を参照することにより、対応する話題評価テーブル 193 を取得し、取得した話題評価テーブル 193 に記録されているト
25 ピックス ID 201 の数をカウントすることによって、処理対象の単語を含む話題の数を取得する。

ステップ S 6 5 において、文書内容処理部 3 3 は、処理対象の単語を含む話題の数が所定数以上であるか否かを判定する。処理対象の単語を含む話題の数が所定数以上であると判定された場合、処理はステップ S 6 6 に進む。ステップ S 6 6 において、文書内容処理部 3 3 は、処理対象の単語を、不要語ベクトル（不要語を構成要素とする）に追加する。これにより、多数の話題に共通して含まれると考えられるあいさつなどの日常的な単語が不要語ベクトルに追加される。

ステップ S 6 7 において、文書内容処理部 3 3 は、不要語である処理対象の単語に対応する記録を削除するため、各話題にそれぞれ対応する話題ファイル 1 6 1、話題単語テーブル 1 8 1、単語インデックステーブル 1 9 1、および話題評価値テーブル 1 9 3 を更新する。この後、処理はステップ S 6 2 に戻り、以降の処理が繰り返される。

なお、ステップ S 6 5 において、処理対象の単語を含む話題の数が所定数よりも小さいと判定された場合にも、ステップ S 6 6 および S 6 7 はスキップされて、処理は、ステップ S 6 2 にリターンする。

その後、ステップ S 6 2 において、ステップ S 4 の処理で生成した単語インデックステーブル 1 9 1 に記録されている単語のうち、以降の処理の対象としていない単語が存在しないと判定された場合、処理はステップ S 6 8 に進む。ステップ S 6 8 において、文書内容処理部 3 3 は、再びステップ S 6 1 の処理と同様に、単語ベクトルが小さい話題、すなわち、対応する単語ベクトル 1 6 7 を構成する単語の数が所定数（例えば、5 個）以下である話題を除去する。これにより、日常的な単語ばかりで構成されているとみなされる話題が除去される。この段階で、話題は特徴的な単語から構成される単語ベクトル 1 6 7 によって象徴されることになる。処理は、図 7 のステップ S 6 にリターンする。

ステップ S 6 において、文書内容処理部 3 3 は、不要語が削除された各単語ベクトル 6 7 を構成する全ての単語について、その出現頻度および複数の文書に亘る分布状況を求め、各話題における評価値を演算する。評価値の演算には、例えば tf・idf 法を用いる。ステップ S 7 において、文書特徴データベース作成部 3

4は、ステップS6で演算した各単語に対する評価値を、次の条件に基づいて修正する。

例えば、送信した電子メールに含まれる単語の評価値がより大きくなるように修正を行う。送信した電子メールに含まれる単語を特定するためには、ステップ

5 S2の処理で生成した各話題に対応する話題ファイル161の連結本文168に挿入した、所定の文字列（例えば”soshin-shuryo”）を検出し、当該所定の文字列以前の単語を、送信した電子メールに含まれる単語として特定すればよい。

また例えば、属する電子メールの数が多い話題に含まれる単語の評価値が、属する電子メールの数に対応して大きくなるように修正を行う。例えば、属する電子メールの数を m とした場合、修正前の評価値に対し、1次関数値 $a \cdot m$ （ a は定数）、対数関数値 $\log(m)$ などの単調増加関数値を乗算する。この修正は、電子メールのような時間的に継続するやりとりでは、以前の文書に登場した単語が、次の文書では指示代名詞によって置換されることが多いので、話題に属する電子メールの数が多くなるほど、単語の評価値が相対的に小さくなってしまう傾向にあることを考慮したものである。

さらに例えば、通信頻度が高い相手と通信した電子メールに含まれる単語、および特定名詞（定義した興味語、一般名、地名、組織名など）などの評価値がより大きくなるように修正を行う。なお、特定名詞に対する評価値の修正方法については、特願2001-379511号として提案した発明を適用することができる。

ステップS8において、文書特徴データベース作成部34は、ステップS6で演算され、ステップS7で修正された各単語に対する評価値を、話題ファイル161および話題単語テーブル181の単語ベクトル167、並びに単語インデックステーブル191の中の話題評価値テーブル193に記録する。これにより、各単語ベクトル167を構成する単語170の全ての要素が決定されたことになる。また、文書特徴データベース作成部34は、各話題にそれぞれ対応する特徴ベクトル169を確定して記録する。さらに、文書特徴データベース作成部13

4は、各単語ベクトル167について、構成する単語を評価値が大きい順に並べ替える。

ステップS9において、文書特徴データベース作成部34は、この段階で残っている話題を再び選抜する。ステップS9の処理、すなわち第2次話題選抜処理5について、図18のフローチャートを参照して説明する。なお、この第2次話題選抜処理は、各話題に対して実行される。

ステップS71において、文書特徴データベース作成部34は、話題に対応する単語ベクトル167を構成する単語のうち、評価値が最大のもの（あるいは、上位の2, 3語）を検出する。ステップS72において、文書特徴データベース作成部34は、ステップS71で検出した単語の評価値が所定値以上であるか否10かを判定する。検出した単語の評価値が所定値以上であると判定された場合、処理はステップS73に進む。

ステップS73において、文書特徴データベース作成部34は、評価値が所定値以上の当該単語を推薦話題候補ベクトルに追加する。ステップS72において、15ステップS71で検出した単語の評価値が所定値よりも小さいと判定された場合、処理はステップS74に進み、文書特徴データベース作成部34は、当該話題を削除する。すなわち、評価値が所定値より小さい単語は興味が少ないものと判断されて検索対象から除外される。

ステップS73またはステップS74の処理の後、すなわち、当該話題に対する20第2次話題選抜処理が終了された後、次の話題に対する第2次話題選抜処理が開始される。全ての話題について第2次話題選抜処理が終了されると、処理は、図7のステップS10にリターンする。

ステップS10において、文書特徴データベース作成部34は、ステップS9の処理で推薦話題候補ベクトルに追加された話題に基づいて、推薦話題を確定す25る。ステップS10における推薦話題確定処理について、図19のフローチャートを参照して説明する。

ステップS 8 1において、文書特徴データベース作成部3 4は、ステップS 9の処理で推薦話題候補ベクトルに追加された要素（各単語ベクトル1 6 7）から、構成する単語の評価値の最大値に注目し、評価値の最大値が大きい順に所定数（例えば、2 0 0）だけ単語ベクトル1 6 7を検出し、それぞれに対応する所定数の話題を取得する。

ステップS 8 2において、文書特徴データベース作成部3 4は、ステップS 8 1の処理で取得された話題が検索条件に合致するか否かを判定し、検索条件に合致すると判定した場合、ステップS 8 3に進む。ここで検索条件とは、例えば、特定期間の話題であるか、特定の相手とやりとりした話題であるか、特定の単語を含む話題であるか、視聴履歴から抽出された単語（例えば、番組名、ジャンル名、あるいは出演者名）を含む話題であるか、または、初期登録された嗜好情報（好みのジャンル名や出演者名）を含む話題であるかなどである。

ステップS 8 3において、文書特徴データベース作成部3 4は、検索条件に合致した話題を推薦話題ベクトルに追加する。ステップS 8 2において、取得された話題が検索条件に合致しないと判定された場合、ステップS 8 3の処理はスキップされる。その後、処理は、図7のステップS 1 1にリターンする。

ステップS 1 1において、文書特徴データベース作成部3 4は、ステップS 1 0の処理で確定された推薦話題ベクトルに基づいて、データベースを作成する。

具体的には、推薦話題ベクトルに追加されている話題の話題ファイル1 6 1（図1 0）をフィルタリングし、後述する番組情報の検索処理において必要となるトピックスID 1 6 2、日時情報1 6 3、サブジェクト情報1 6 4、メンバー情報1 6 5、および単語ベクトル1 6 7（興味語ベクトル2 1 2）を抽出して、図2 0に示されるような興味データ2 1 1を作成する。そして、文書特徴データベース作成部3 4は、興味データ2 1 1を、当該ユーザのユーザID、メールアドレス、ログインアカウント、またはパスワードなどに対応付けてデータベースを作成し、記憶部5 9に記録する。ただし、ステップS 1 1の処理は、ステッ

プ S 1 0 までの一連の処理に継続して実行される場合と、一連の処理に継続せず、所定のタイミングにおいて実行される場合がある。

なお、興味データ 2 1 1 は、番組情報を検索するためのキーワードとなるため、単語ベクトル 1 6 7 を、新たに興味語ベクトル 2 1 2 として定義する。また単語
5 ベクトル 1 6 7 を構成する単語（興味語）のうち、評価値が最大のもの（あるいは、上位の 2, 3 語）のみを興味語ベクトル 2 1 2 として用いるようにしてもよい。また、興味データ 2 1 1 には、例えば、当該話題の電子メールの重要語を含む 1 文書を保持させるようにしてもよい。

ところで、テレビ番組の EPG データでよく使用される単語、例えば、「ミュージック」、「ニュース」、または、出演者のニックネームなどの未定義語（表
10 記ゆれや短縮語を含む）は、例えば、“名詞ーテレビ語”といったように、単なる一般名詞とは異なるように、予め定義しておく。そして、形態素解析の結果、“名詞ーテレビ語”に分類された単語は、図 7 のステップ S 6 における評価値の演算時に、所定の重みがかけられ、“名詞ーテレビ語”の単語の評価値がより大
15 きくなるように修正される。このようにしておくことにより、電子メールに含まれる単語の中で“名詞ーテレビ語”の単語が、興味語ベクトル 2 1 2 を構成する単語に含まれる（すなわち、興味語になる）可能性が高まる。

さらに、複数のユーザが、1 つのメール 1 2 を共用している場合、メールアカウント毎に、電子メールの形態素解析が施され、ユーザ毎に興味データ 2 1 1 が
20 作成される。データベースには、メールアカウントをキーにして、ユーザ毎の興味データ 2 1 1 が記録される。

以上のデータベース作成処理が実行されることにより、送受信した電子メールの文書から抽出された興味語で構成される興味データ 2 1 1 がデータベース内に蓄積されることになる。

25 また、データベース作成処理をユーザが強制的に中断することができるようにするために、中断要求があった場合、中断された時点で処理済みの文書を記録し、再開要求があった場合、未処理の文書から処理を再開するようにしてもよい。

なお、データベース作成処理は、エージェントプログラム 1 1 が実行された場合に開始されるものとしたが、任意のタイミングで開始させることも可能である。さらに、このようにして作成されたデータベースは、所定の条件が満たされたときに更新される。

- 5 次に、蓄積部 2 1 によるデータベースの更新のタイミングについて説明する。データベースは、上述したデータベース作成処理によって作成されるが、次のような第 1 乃至第 3 の状況になった場合、データベースが更新される。

すなわち、第 1 の状況として、データベースが作成または更新されてから所定の期間が経過している場合、データベース内の関連情報が古くなってしまうので
10 更新が行われる。

第 2 の状況として、データベースに蓄積されている関連情報のうちの所定の割合が提示済みとなった場合、データベース内の同じ関連情報が繰り返し提示されたり、提示する関連情報が不足したりしてしまうので更新が行われる。

第 3 の状況として、特徴抽出に用いた文書が電子メールである場合、電子メールの送受信が繰り返されていると、その文書の内容が変化するので更新が行われ
15 る。

なお、データベースの更新が必要である状況になった場合（例えば、イベント管理部 4 1 がタイマ 4 1 A を監視し、所定の期間が経過したとき）、ユーザに対して更新を指示するように促すこともできるし、ユーザに対する更新指示の促し
20 を実行することなく、自動的にデータベースを実行するように設定することも可能である。また、ユーザが指定する任意のタイミングで更新することも勿論可能である。

これら第 1 乃至第 3 の状況を考慮したデータベース更新処理について、図 2 1 のフローチャートを参照して説明する。このデータベース更新処理は、エージェントプログラム 1 1 が実行する処理のうちの 1 つであり、エージェントプログラム 1 1 の起動とともに開始され、エージェントプログラム 1 1 が終了されるまで
25

繰り返し実行される。なお、この処理が開始される以前において、既に上述したデータベース作成処理が実行されており、データベースが存在するものとする。

ステップS 9 1において、エージェントプログラム1 1の蓄積部2 1は、作成済みのデータベースの更新が必要であるか否かを判定し、更新が必要であると判定されるまで待機する。この判断基準は、例えば図2 2に示すようなユーザインタフェースの画面を用いて予めユーザが設定するものとする。図2 2の例では、4つの条件が示されており、ユーザによって左端の口印（チェックボックス）がチェックされた場合、対応する条件が有効となる。なお、1番目の条件では回数が設定可能とされており、3番目の条件では日数が設定可能とされている。

- 10 ステップS 9 1において、更新が必要であると判定された場合、処理は、ステップS 9 2に進む。ステップS 9 2において、蓄積部2 1は、データベースを自動的に更新するように設定されているか否かを判定し、自動的に更新するように設定されていないと判定した場合、ステップS 9 3に進む。一方、ステップS 9 2において、自動的に更新するように設定されていると判定された場合、ステップS 9 3の処理はスキップされる。

- 15 ステップS 9 3において、エージェントプログラム1 1の提示部2 2は、データベースの更新が必要である旨をユーザに通知するとともに、さらに、その通知に対応して、ユーザから更新の指示がなされたか否かを判定する。ユーザから更新の指示がなされたと判定された場合、処理はステップS 9 4に進む。反対に、
- 20 ユーザから更新の指示がなされないと判定された場合、処理はステップS 9 1に戻り、以降の処理が繰り返し実行される。

- 25 ステップS 9 4において、エージェントプログラム1 1の蓄積部2 1は、データベースを更新する。具体的には、文書取得部3 1乃至文書内容処理部3 3が、電子メールの電子メールボックスファイル（特定の拡張子mbx等が付与されていることが多い）を検出し、その更新日時を取得して、以前に取得した更新日時と比較し、異なる日付と異なるファイルサイズであれば、ファイルが更新されていると判断し、追加または変更された部分を抽出する。この場合、電子メールの

グループ化、ヘッダの解析、形態素解析、特徴ベクトル算出等、一連のファイル内の分析が行われ、得られる重要語が関連情報検索部 35 に供給される。

- ただし、メールグループ（話題）が変化せず（所定的话题に新たに追加された電子メールがなく）、分析の結果、更新以前の重要語（検索用キーワード）と更新後の重要語が同じであれば、評価値等の計算値だけを変更し、関連情報検索部 35 による関連情報の検索を実行しないようにしてもよい。

- あるいは、全ての電子メールグループが変化せずに一定期間が経過した場合、グループの特徴ベクトルのうち、前回、評価値が 1 番目と 2 番目の単語を検索語としていたものを、例えば評価値が 3 番目と 4 番目の単語を検索語に変更して検索し、検索結果を取得するようにしてもよい。

以上説明したように、データベース更新処理においては、追加された文書や変更された文書だけを処理の対象とするので、データベース作成処理を繰り返し実行する場合に比べて、処理時間が短縮される。

- 次に、サーバ 6 のサーバプログラム 101 によるデータベース作成処理について、図 23 のフローチャートを参照して説明する。このデータベース作成処理は、サーバプログラム 101 が実行する処理のうちの 1 つであり、サーバプログラム 101 が起動された状態において、データベースが未だ作成されていないときに開始される。

- ステップ S101 において、番組メタデータ取得部 121 は、データベース作成の素として分析する EPG データなどの番組情報（番組メタデータ）を取得し、データ内容処理部 122 に供給する。

- 図 24 は、ステップ S101 において取得される番組メタデータ 220 の一例を示している。番組メタデータ 220 は、当該番組の番組名を示すタイトル 221、当該番組の分類（例えば、ドラマ、映画、ニュース、スポーツ、あるいは音楽など）を示すジャンル 222、当該番組の放送時間（例えば、朝、昼、夕方、ゴールデン、あるいは深夜など）を示す時間帯情報 223、当該番組が放送されるチャンネル（例えば、NHK 総合、日本テレビ、あるいは TBS（いずれも商

標) など) を示す放送局 2 2 4、当該番組に出演している人物を示す出演者情報 2 2 5、当該番組の脚本、原作、あるいは演出した人物を示す脚本・原作・演出情報 2 2 6、および、当該番組のあらすじや見所などを示す内容 (キーワード) 情報 2 2 7 から構成される。

- 5 図 2 3 に戻る。ステップ S 1 0 2 において、データ内容処理部 1 2 2 は、ステップ S 1 0 1 の処理で取得された番組メタデータ 2 2 0 に形態素解析を施し、番組情報 (番組名、ジャンル名、放送局名、時間帯情報、出演者名、およびキーワードなど) を抽出する。ステップ S 1 0 3 において、データベース作成部 1 2 3 は、データ内容処理部 1 2 2 によって抽出された番組情報を番組毎にデータベース化して、記憶部 1 4 7 に記録する。
- 10

以上のデータベース作成処理が実行されることにより、番組メタデータ 2 2 0 から抽出された番組情報が推薦番組用データベース内に蓄積されることになる。また、データベース作成処理をユーザが強制的に中断することができるようにするために、中断要求があった場合、中断された時点で処理済みの文書を記録し、

15 再開要求があった場合、未処理の文書から処理を再開するようにしてもよい。

なお、データベース作成処理は、サーバプログラム 1 0 1 が実行された場合に開始されるものとしたが、任意のタイミングで開始させることができるとともに、所定の期間毎 (例えば、EPG データ更新毎) に更新される。

- 次に、図 7 の処理でパーソナルコンピュータ 1 のデータベースに記録された興味データ 2 1 1 にマッチングする番組情報を、図 2 3 の処理で作成されたサーバ 20 6 の推薦用番組データベースから検索する処理について、図 2 5 のフローチャートを参照して説明する。

- ステップ S 1 2 1 において、HDD レコーダ 2 の CPU 7 1 は、ユーザによって入力部 7 6 が操作され、ログインされたか否かを判定し、ログインされたと判定されるまで待機する。ステップ S 1 2 1 において、ユーザによりログインされたと判定された場合、ステップ S 1 2 2 に進み、CPU 7 1 は、興味データを取得する
- 25

ためのコマンドとともにユーザのログイン情報を、通信部 77 を介してパーソナルコンピュータ 1 に送信する。

ステップ S 1 1 1 において、パーソナルコンピュータ 1 のデータベース問い合わせ部 42 は、HDD レコーダ 2 から送信されてきた興味データ取得コマンドを受信し、文書特徴データベース作成部 34 によって作成されたデータベースを検索し、ログイン情報（ログインアカウントやパスワードなど）に対応する、興味データ 2 1 1（図 20）を読み出し、通信部 60 を介して HDD レコーダ 2 に送信する。

ステップ S 1 2 3 において、HDD レコーダ 2 の CPU 71 は、ログインされたユーザに対応する興味データ 2 1 1 を、パーソナルコンピュータ 1 から受信し、RAM 73 に記録させる。ステップ S 1 2 4 において、CPU 71 は、受信した興味データ 2 1 1 を、通信部 77 およびネットワーク 5 を介してサーバ 6 に送信し、興味データ 2 1 1 にマッチする番組情報の検索を要求する。

ステップ S 1 3 1 において、サーバ 6 のイベント管理部 131 は、ネットワーク 5 を介して HDD レコーダ 2 から送信されてきた興味データ 2 1 1 を受信する。そして、イベント管理部 131 は、興味データ 2 1 1 をデータベース問い合わせ部 132 に供給して検索要求を通知する。ステップ S 1 3 2 において、データベース問い合わせ部 132 は、イベント管理部 131 からの検索要求の通知に対応して、データベース作成部 123 によって作成された推薦番組用データベースを検索し、検索要求に含まれる興味データ 2 1 1 にマッチする番組情報を推薦用番組として選択する。

ここで、興味データ 2 1 1 を構成する興味語ベクトル 2 1 2 には、出演者のニックネームなどの未定義語、表記ゆれや短縮語、あるいは、言い回し（ニュアンス）の違いなどの単語が含まれている場合がある。従って、サーバ 6 では、それらの単語にマッチする番組情報を検索することができるようになるため、未定義語、表記ゆれ、および短縮語などの対応辞書を持つようにすることが好ましい。

ステップ S 1 3 3 において、番組情報出力部 1 3 3 は、データベース問い合わせ部 1 3 2 で選択された推薦用番組（番組情報）を検索結果として、通信部 1 4 8 およびネットワーク 5 を介して HDD レコーダ 2 に送信する。

5 ステップ S 1 2 5 において、HDD レコーダ 2 の CPU 7 1 は、ネットワーク 5 を介してサーバ 6 から送信されてきた番組情報に含まれる録画開始時刻、録画終了時刻、およびチャンネルなどに基づいて、RAM 7 3 に番組の録画予約を設定するとともに、チューナ 7 9、復調部 8 0、デコーダ 8 1、および HDD 7 8 などを制御する。これにより、録画予約の開始時刻になると、RAM 7 3 に設定されている番組情報が読み出され、録画予約が実行される。

10 以上のようにして、ユーザの興味にマッチした番組情報が検索され、HDD レコーダ 2 に録画予約が自動的に設定される。すなわち、ユーザがやりとりした電子メールから抽出した興味語を基に番組情報を検索するため、日常の興味が反映されたタイムリーな番組推薦が可能となる。

15 またステップ S 1 1 1 の処理でデータベースから興味データ 2 1 1 を読み出して送信する際に、興味データ 2 1 1 の興味語ベクトル 2 1 2 に含まれる各単語（興味語）のうち、評価値の時系列推移に着目して、所定条件を満たす単語（重要語）のみを選択するようにしてもよい。ここで所定の選択条件としては、例えば、条件 1 を「現時点以前の所定の期間 Y（例えば、5 週間）において、異なる 2 以上の話題で、当該単語の評価値が所定の閾値 B 以上であること」、条件 2 を
20 「条件 1 における異なる 2 以上の話題のうち、最も古い話題と最も新しい話題が所定の期間 Z 以上離れていること」などとする。

このような条件を用いることにより、ユーザが高い関心を持っていると思われる単語（重要語）や、ユーザが現時点で意識していない意外性があると思われる単語にマッチした番組情報を推薦することが可能となる。

25 なお、上述した例においては、電子メールの形態素解析、およびその解析結果から抽出される興味データ 2 1 1 の作成処理をパーソナルコンピュータ 1 で行い、HDD レコーダ 2 には、興味データ 2 1 1 のみを送信させるようにしたが、HDD レ

コード 2 に、パーソナルコンピュータ 1 のエージェントプログラム 1 1 と同等の機能を有するプログラムを実装させるようにすれば、電子メールの形態素解析や興味データ 2 1 1 の作成処理を HDD レコーダ 2 に行わせることも可能である。

5 また、パーソナルコンピュータ 1 では、電子メールの形態素解析、および解析結果に基づいて興味データ 2 1 1 を作成し、HDD レコーダ 2 では、パーソナルコンピュータ 1 で作成された興味データ 2 1 1 を取得してサーバ 6 に送信し、興味データ 2 1 1 にマッチする番組情報の推薦を受けるようにし、サーバ 6 では、興味データ 2 1 1 にマッチする番組情報を検索し、検索結果を HDD レコーダ 2 に送信するようにして、処理の分散を図るようにした。

10 本発明は、これに限らず、例えば、パーソナルコンピュータ 1 が、サーバ 6 で作成された推薦番組用データベースの番組情報を一括して提供してもらい、番組情報の検索を行うこともできる。その場合、HDD レコーダ 2 は、パーソナルコンピュータ 1 から供給される番組情報に含まれる録画開始時刻、録画終了時刻、およびチャンネルなどを取得し、録画予約を実行する。

15 さらに、サーバ 6 が、例えば、1 つのインターネットサービスプロバイダであり、上述した推薦番組の配信サービスやメールサービスを行っている場合、ユーザのメールサーバに直接アクセスして、サーバ 6 側で電子メールの形態素解析、およびその解析結果に基づいて興味データ 2 1 1 を作成するようにしてもよい。

ところで、上述した処理では、ユーザに気付かれずに、ユーザの興味にマッチした番組が検索され、自動的にその番組の録画予約が実行されてしまう。そこで、例えば、HDD レコーダ 2 の電源オン時、録画予約確認時、あるいは、録画番組の再生時に、録画予約された番組の推薦理由を提示させるようにする。その場合の処理について、図 2 6 のフローチャートを参照して説明する。

20 ステップ S 1 4 1 において、HDD レコーダ 2 の CPU 7 1 は、提示条件を満たすか否かを判定し、提示条件を満たすと判定されるまで待機する。ここで提示条件とは、例えば、HDD レコーダ 2 に番組の録画予約が設定され、かつ、テレビジョン受像機 3 が電源オン状態であるか、あるいは、HDD レコーダ 2 に番組の録画予

約が設定され、かつ、HDD レコーダ 2 の GUI 画面が表示可能状態であるかなどである。

ステップ S 1 4 1 において、提示条件を満たすと判定された場合、ステップ S 1 4 2 に進み、CPU 7 1 は、図 2 5 のステップ S 1 2 3 の処理で RAM 7 3 に記録された興味データ 2 1 1 に含まれるトピックス I D 1 6 2 乃至メンバー情報 1 6 5 (図 2 0) に基づいて、推薦理由 (背景情報) をテレビジョン受像機 3 に出力し、表示させる。

図 2 7 は、推薦理由の表示例を示している。同図に示されるように、エージェント 2 3 1 を登場させるとともに、エージェント 2 3 1 の台詞として吹き出し 2 3 2、および、ユーザにより操作可能な入力ウィンドウ 2 3 3 が表示される。吹き出し 2 3 2 の中には、例えば「2 0 0 1 年 4 月 9 日、太郎さんと「ワインありがとう」ってメールしたよね。このまま録画する？」の台詞が表示される。

また、吹き出し 2 3 2 の表示と同期して、吹き出し 2 3 2 に表示された台詞と同じ意味を持つ言語 (例えば、日本語や英語など) の音声信号が音声合成部 (図 15 示せず) によって合成されて出力させるようにすることができる。ただし、吹き出し 2 3 2 の表示の有無や台詞に対応する音声の出力の有無はエージェントプログラム 1 1 が適宜設定するか、ユーザが任意に設定できるようにすることが可能である。

また入力ウィンドウ 2 3 3 には、番組を録画予約するときに選択する「録画する」ボタン、番組の録画予約をキャンセルするときに選択する「キャンセル」ボタンが表示される。

ユーザは、吹き出し 2 3 2 に表示された台詞によって、「太郎さん」とメールした内容 (いまの場合、「ワインありがとう」) から、「△△△のワイン紀行」の番組がサーバ 6 から推薦され、HDD レコーダ 2 に録画予約されたことを知ることができる。そして、ユーザは、このまま録画を開始させる場合には、「録画する」ボタンを選択し、録画をキャンセルさせる場合には、「キャンセル」ボタンを選択する。なお、ユーザによりいずれのボタンも選択されなかった場合を想定

して、タイムアウトをトリガにして、いずれかのボタンが自動的に選択されるようにしてもよい。

図26に戻る。ステップS143において、CPU71は、ユーザによって入力部76が操作されて「キャンセル」ボタンが選択されたか否かを判定し、「キャンセル」ボタンが選択されたと判定した場合、ステップS144に進み、RAM73に設定されている録画予約を削除させる。

ステップS143において、「キャンセル」ボタンが選択されなかった、すなわち、「録画する」ボタンが選択されたと判定された場合、ステップS144の処理はスキップされる。

10 以上のようにして、自動的に録画予約された番組の推薦理由を表示させることができ、その番組を録画するか否かをユーザに決定させることができる。

また上述した例では、HDDレコーダ2からテレビジョン受像機3に推薦理由を表示させるようにしたが、これに限らず、例えば、図25のステップS133の処理において、サーバ6がHDDレコーダ2に検索結果を送信する際に、パーソナルコンピュータ1にも検索結果を送信するようにし、パーソナルコンピュータ1の電源オン時、あるいは、エージェントプログラム11の起動時に、推薦理由を表示させることもできる。その場合の推薦理由の表示例を図28に示す。

図28の例の場合、吹き出し232の中には、例えば「2001年4月9日、太郎さんとメールした「ワインありがとう」から「△△△のワイン紀行」という番組の推薦があったよ。録画する？」の台詞が表示される。

ユーザは、吹き出し232に表示された台詞によって、「太郎さん」とメールした内容（いまの場合、「ワインありがとう」）から、「△△△のワイン紀行」の番組がサーバ6から推薦され、HDDレコーダ2に録画予約されたことを知ることができる。そして、ユーザは、このまま録画を開始させる場合には、「録画する」ボタンを選択し、録画をキャンセルさせる場合には、「キャンセル」ボタンを選択する。「キャンセル」ボタンが選択された場合、パーソナルコンピュータ

1のCPU41は、HDDレコーダ2に、録画予約の削除コマンドを送信し、録画予約の削除を実行させる。

以上のようにして、自動的に録画予約された番組の推薦理由を、HDDレコーダ2を介してテレビジョン受像機3に、またはパーソナルコンピュータ1に表示させることができる。また推薦理由には、ユーザが電子メールでやりとりした話題、日時、やりとりした相手、またはサブジェクトなどが含まれているため、単に「××番組が推薦されました」という提示に較べ、より興味を引くことができる。

以上においては、ユーザがやりとりした電子メールや作成した文書などに含まれる単語（興味語）にマッチした番組情報を検索し、推薦するようにしたが、これに限らず、例えば、ユーザが初期登録した嗜好情報（好みのジャンル名や出演者名など）、または、ユーザが視聴した番組の視聴履歴などを分析し、その分析結果から興味語を抽出して、マッチングする番組情報を検索することもできる。その場合の処理について、図29のフローチャートを参照して説明する。

ステップS151において、パーソナルコンピュータ1のイベント管理部41は、内蔵するタイマ41Aを参照することによって時間の経過を監視し、所定期間が経過したか否かを判定し、所定期間が経過するまで待機する。そして、所定期間が経過したと判定された場合、イベント管理部41は、イベント発生をデータベース問い合わせ部42に通知し、ステップS152に進む。

ステップS152において、データベース問い合わせ部42は、文書特徴データベース作成部34によって作成されたデータベースから、初期登録された嗜好情報または視聴履歴が形態素解析されて記録された興味データ211を読み出し、ネットワーク5を介してサーバ6に送信する。

ステップS171において、サーバ6のイベント管理部131は、ネットワーク5を介してパーソナルコンピュータ1から興味データ211を受信する。そして、イベント管理部131は、興味データ211をデータベース問い合わせ部132に供給して検索要求を通知する。ステップS172において、データベース問い合わせ部132は、イベント管理部131からの検索要求の通知に対応して、

データベース作成部 1 2 3 によって作成された推薦番組用データベースを検索し、検索要求に含まれる興味データにマッチする番組情報を推薦用番組として選択する。

5 ステップ S 1 7 3 において、番組情報出力部 1 3 3 は、データベース問い合わせ部 1 3 2 で選択された推薦用番組（番組情報）を検索結果として、通信部 1 4 8 およびネットワーク 5 を介してパーソナルコンピュータ 1 および HDD レコーダ 2 に送信する。

10 ステップ S 1 6 1 において、HDD レコーダ 2 の CPU 7 1 は、ネットワーク 5 を介してサーバ 6 から送信されてきた番組情報に含まれる録画開始時刻、録画終了時刻、およびチャンネルなどに基づいて、RAM 7 3 に番組の録画予約を設定するとともに、チューナ 7 9、復調部 8 0、デコーダ 8 1、および HDD 7 8などを制御する。これにより、録画予約の開始時刻になると、RAM 7 3 に設定されている番組情報が読み出されて、録画予約が実行される。

15 ステップ S 1 5 3 において、イベント管理部 4 1 は、検索結果を受信し、イベント発生をデータベース問い合わせ部 3 2 に通知する。データベース問い合わせ部 4 2 は、イベント管理部 4 1 からのイベント発生の通知に対応して、イベント発生に対応する検索結果（番組情報）を取得し、形態素解析を施して単語（特徴語）を抽出し、各単語の評価値を演算する。これにより、検索結果（番組情報）の特徴ベクトルが算出される。

20 ステップ S 1 5 4 において、データベース問い合わせ部 4 2 は、文書特徴データベース作成部 3 4 が作成したデータベースを検索し、ステップ S 1 5 3 の処理で算出された特徴ベクトルと、データベースに記録されている話題毎の特徴ベクトルとの内積を両者の類似度として算出し、類似度が所定の条件（例えば、類似度が最大、もしくは類似度が所定の閾値以上）を満たす話題を抽出する。

25 このとき、例えば、特定のジャンル名（ユーザによって初期登録された番組のジャンル名）を含む話題を予め選抜しておくようにして、検索結果（番組情報）

に類似する話題を効率良く抽出することができるようにしてもよい。また、一般的な番組のジャンル名を含む話題を一通り選抜しておくようにしてもよい。

ステップS 1 5 5において、データベース問い合わせ部4 2は、ステップS 1 5 4の処理で抽出された話題を構成する文書のうち、最新の文書を選択し、イベント管理部4 1を介して、または直接的に、関連情報提示部4 3に供給する。ステップS 1 5 6において、エージェント制御部2 3は、ステップS 1 5 5の処理で選択した文書の属性情報を、選択（推薦）した理由としてデスクトップに表示させる（図2 8）。

ステップS 1 5 7において、エージェントプログラム1 1は、入力部5 6が操作されて「キャンセル」ボタンが選択されたか否かを判定し、「キャンセル」ボタンが選択されたと判定された場合、ステップS 1 5 8に進み、録画予約の削除コマンドをHDDレコーダ2に送信する。ステップS 1 5 7において、「キャンセル」ボタンが選択されなかった、すなわち、「録画する」ボタンが選択されたと判定された場合、ステップS 1 5 8の処理はスキップされる。

ステップS 1 6 2において、HDDレコーダ2のCPU7 1は、パーソナルコンピュータ1から送信されてきた録画予約の削除コマンドを受信し、ステップS 1 6 3において、RAM7 3に設定されている録画予約を削除させる。

以上のようにして、ユーザによって初期登録された嗜好情報や視聴履歴などの多くの情報を形態素解析して、有意な興味データ2 1 1を作成するため、その興味データ2 1 1にマッチした、意外性のある番組情報を推薦することができる。また、推薦番組に類似する話題（または文書）を文書特徴データベース作成部3 4が作成したデータベースから逆引きすることで、推薦された番組に関連する推薦理由を取得することができる。従って、ユーザは、個人的な結びつきによって番組が推薦されたかのように感じることもできる。

以上においては、ユーザがやりとりした電子メールや作成した文書、あるいは、嗜好情報や視聴履歴から興味データ2 1 1を作成し、それにマッチした番組情報を検索し、推薦するようにした。すなわち、パーソナルコンピュータ1または

HDD レコーダ 2 は、予め作成された興味データ 2 1 1 をデータベースから読み出し、ネットワーク 5 を介してサーバ 6 に送信することによって、マッチングする番組の推薦を受けるようにした。従って、ユーザの意思に関わらず、ユーザの興味があると思われる番組を自動的に検索し、推薦するというものである。

- 5 ところで、当然のことながら、ユーザの意思によって、番組の録画予約を行うようにすることもできる。次に、図 3 0 のフローチャートを参照して、ユーザが、例えば、外出先において、所有しているデジタル携帯電話機 4 を用いて、番組の録画予約を行う場合の処理について説明する。

- ユーザは、デジタル携帯電話機 4 を操作し、図 3 1 に示されるような、番組の
10 録画予約をするための電子メール（以下、予約メールと記述する）2 4 1 を作成する。同図に示されるように、予約メール 2 4 1 の本文には、「今日、音楽番組を録画して。特にクラシックものは絶対。あればジャズ番組も。あとワールドカップ情報もお願い。」といったメッセージが記載されている。

- すなわち、ユーザは、予約メール 2 4 1 を作成するにあたって、記述フォーマットを気にする必要なく、自由文で作成すればよく、予約メール 2 4 1 には、録
15 画予約したい番組の番組名や出演者名などの一部、あるいは、ジャンル名を入力するだけでもよい。ここで、予約メール 2 4 1 の受信側において、その電子メールが番組を録画予約するための電子メール（すなわち、予約メール 2 4 1）であることを認識することができるように、例えば、サブジェクト（表題）に、「予
20 約メール」あるいは「録画予約」といった定型文（図 3 1 の例の場合、「予約メール」）を入力するようにする。

- ユーザは、作成した予約メール（図 3 1）を、自宅にあるパーソナルコンピュータ 1 に送信するように、デジタル携帯電話機 4 を操作する。これにより、ステ
ップ S 1 8 1 において、デジタル携帯電話機 4 は、ユーザの操作に対応する入力
25 信号を受けて、作成された予約メール 2 4 1 を、ネットワーク 5 を介してパーソナルコンピュータ 1 に送信する。

ステップS 1 9 1において、パーソナルコンピュータ 1 のイベント管理部 4 1 は、ネットワーク 5 を介して予約メール 2 4 1 を受信したことを検知して、データベース問い合わせ部 4 2 に通知する。ステップS 1 9 2において、データベース問い合わせ部 4 2 は、イベント管理部 4 1 からのイベント発生に対応する予約メール 2 4 1 を取得する。このときデータベース問い合わせ部 4 2 は、予約メール 2 4 1 のサブジェクトに「予約メール」と記述されていることにより、番組を録画予約するための電子メールであると認識する。そしてデータベース問い合わせ部 4 2 は、取得した予約メール 2 4 1 に形態素解析を施して単語を抽出し、不要語を除外して興味語ベクトル（特徴ベクトル）を作成（算出）する。

10 いまの場合、予約メール 2 4 1 に記載された「今日、音楽番組を録画して。特にクラシックものは絶対。あればジャズ番組も。あとワールドカップ情報もお願い。」から、形態素解析によって、「今日、音楽、クラシック、ジャズ、ワールドカップ」が興味語ベクトルとして抽出される。また、「今日」はこの時点で日時情報（例えば、2 0 0 1 年 4 月 9 日）に変換される。

15 なお、カテゴリに木構造を予め定義しておくことによって、形態素解析の分析精度を向上させることができる。例えば、音楽のカテゴリに{クラシック、ジャズ、ポップス、ロック、演歌、・・・}などと定義しておき、形態素解析によって抽出された興味語ベクトルに適用させると、「クラシック、ジャズ、ワールドカップ」の 3 つの単語（興味語）で興味語ベクトルが構成される。

20 ステップS 1 9 3において、データベース問い合わせ部 4 2 は、ステップS 1 9 2 の処理で作成された興味語ベクトルを、ネットワーク 5 を介してサーバ 6 に送信する。

25 ステップS 2 1 1において、サーバ 6 のイベント管理部 1 3 1 は、ネットワーク 5 を介してパーソナルコンピュータ 1 から興味語ベクトルを受信する。そして、イベント管理部 1 3 1 は、興味語ベクトルをデータベース問い合わせ部 1 3 2 に供給して検索要求を通知する。ステップS 2 1 2において、データベース問い合わせ部 1 3 2 は、イベント管理部 1 3 1 からの検索要求の通知に対応して、デー

データベース作成部 1 2 3 によって作成された推薦番組用データベースを検索し、検索要求に含まれる興味語ベクトルにマッチする番組情報を録画予約用の番組として選択する。

5 ステップ S 2 1 3 において、番組情報出力部 1 3 3 は、データベース問い合わせ部 1 3 2 で選択された録画予約用の番組（番組情報）を検索結果として、ネットワーク 5 を介して HDD レコーダ 2 に送信する。

10 ステップ S 2 0 1 において、HDD レコーダ 2 の CPU 7 1 は、ネットワーク 5 を介してサーバ 6 から送信されてきた番組情報に含まれる録画開始時刻、録画終了時刻、およびチャンネルなどに基づいて、RAM 7 3 に番組の録画予約を設定するとともに、チューナ 7 9、復調部 8 0、デコーダ 8 1、および HDD 7 8 などを制御する。これにより、録画予約の開始時刻になると、RAM 7 3 に設定されている番組情報が読み出され、録画予約が実行される。

15 ステップ S 2 0 2 において、CPU 7 1 は、録画予約の設定が完了すると、例えば、図 3 2 に示されるような、番組の録画予約を完了した旨を通知するための電子メール（以下、予約完了メールと記述する） 2 5 1 を作成する。同図に示されるように、予約完了メール 2 5 1 の本文には、「下記の番組を録画しました」といったメッセージとともに、第 1 番目のエントリに、「4CH」の放送局の

20 「19:00~20:00」の時間に「ワールドカップハイライト」が録画予約された旨が記載され、第 2 番目のエントリに、「3CH」の放送局の「21:00~21:54」の時間に「×××クラシック」が録画予約された旨が記載される。

そして、作成された予約完了メール 2 5 1 は、ネットワーク 5 を介してデジタル携帯電話機 4 に送信される。

25 ステップ S 1 8 2 において、デジタル携帯電話機 4 は、ネットワーク 5 を介して HDD レコーダ 2 より送信されてきた予約完了メール 2 5 1 を受信する。そして、ユーザが電子メールを表示させるように操作すると、ステップ S 1 8 3 において、デジタル携帯電話機 4 は、ユーザの操作に対応する入力信号に基づいて、受信した予約完了メール 2 5 1 をディスプレイに表示させる。

以上のようにして、パーソナルコンピュータ 1 のエージェントプログラム 1 1 は、送信されてきた予約メール 2 4 1 を形態素解析し、興味語ベクトル（録画予約に必要な情報）を作成してサーバ 6 に送信する。サーバ 6 は、パーソナルコンピュータ 1 から送信されてきた興味語ベクトルにマッチする所定の番組情報を検索し、HDD レコーダ 2 に送信する。これにより、HDD レコーダ 2 には、自動的に番組の録画予約が行われる。

従って、ユーザは、外出先においても、記述形式を気にすることなく自由文で作成した予約メール 2 4 1 を自宅のパーソナルコンピュータ 1 に送信するだけで、簡単に番組の録画予約を行うことができるとともに、番組の録画予約が完了したことを簡単に知ることができる。

以上説明した本実施の形態においては、ユーザがやりとりした電子メールの内容、ユーザによって初期登録された嗜好情報、および視聴履歴などを形態素解析して興味データ 2 1 1 を作成し、その興味データ 2 1 1 にマッチする番組情報の推薦を受けるようにするか、あるいは、ユーザが作成した予約メール 2 4 1 （番組を録画予約するための電子メール）を形態素解析して興味語ベクトルを作成し、その興味語ベクトルにマッチする番組情報を録画するようにした。従って、ユーザは、日常的な関心からの番組や潜在的な関心からの番組といったように、場面や目的に応じた多様な番組推薦を受けることができる。

また、サーバ 6 は、強調フィルタリング方式を用いて、他のユーザの興味データ 2 1 1 を利用して番組推薦を行うこともできる。その場合、パーソナルコンピュータ 1 または HDD レコーダ 2 は、興味データ 2 1 1 に含まれる情報のうち、プライバシー上問題の無いデータになるようにフィルタリングした後（例えば、番組名は選択せずに、ジャンル名のみを選択した後）、サーバ 6 に送信するようにする。このフィルタリングのレベルは、ユーザが任意に設定することができる。

さらに、図 2 7 の表示例を用いて説明したように、推薦理由の提示時、推薦された番組を録画予約するか否かをユーザにより決定させるため、HDD レコーダ 2 は、録画予約が実行された番組と、録画予約がキャンセルされた番組に分類して、

推薦番組のアクセス履歴を保存することができる。従って、HDD レコーダ 2 は、推薦番組を録画予約する頻度をカウントしたり、あるいは、アクセス履歴に形態素解析を施して新たな興味を抽出することにより、例えば、図 3 3 に示されるような嗜好データ 2 6 0 を作成することが可能になる。

- 5 図 3 3 の例の場合、嗜好データ 2 6 0 は、推薦番組の分類（例えば、ドラマ、映画、ニュース、スポーツ、あるいは音楽など）を示すジャンル 2 6 1、当該推薦番組の番組名を示すタイトル 2 6 2、当該推薦番組の放送時間（例えば、朝、昼、夕方、ゴールデン、あるいは深夜など）を示す時間帯情報 2 6 3、当該推薦番組に出演している人物を示す出演者情報 2 6 4、当該推薦番組のあらすじや見
- 10 所などを示す内容（キーワード）情報 2 6 5、および、当該推薦番組と一緒に視聴していたユーザに関する同席者情報 2 6 6 から構成される。

- なお、同席者情報 2 6 6 の作成方法は、例えば、ユーザが入力部 5 7 を用いて入力するようにしてもよいし、CPU 4 1 が、各ユーザが所有している、デジタル携帯電話機や時計などから常時発信される機器 ID を検知し、機器 ID からユー
- 15 ザを特定して入力するようにしてもよいし、あるいは、各ユーザの会話から音声認識し、認識結果からユーザを特定して入力するようにしてもよい。

- このようにして作成された嗜好データ 2 6 0 は、ユーザが任意にプライバシー上問題の無いデータになるようにフィルタリングされてサーバ 6 に送信される。そして、嗜好データ 2 6 0 を受信したサーバ 6 は、その嗜好データ 2 6 0 にマッ
- 20 チする番組情報を検索し、ユーザが新たに興味を引くような番組を推薦することができる。

- なお、図 2 0 に示した興味データ 2 1 1、および図 3 3 に示した嗜好データ 2 6 0 は、所定のスキーマを定義し、例えば、拡張性のある XML (eXtensible Markup Language) など記述することによって、HTTP (Hypertext Transport
- 25 Protocol) を用いて、任意の場所および任意の機器から、容易に、アクセスすることができるようにしてもよい。

また、パーソナルコンピュータ 1 と HDD レコーダ 2 は、Ethernet (R) を介してデータの授受が行われるものとして説明したが、例えば、i.Link (登録商標)、または IEEE (Institute of Electrical and Electronics Engineers) 802.11a, IEEE 802.11b, Bluetooth (登録商標) などの無線 LAN を介してデータの授受を行うようにしてもよい。さらに、磁気ディスク、光ディスク、光磁気ディスク、もしくは半導体メモリよりなるリムーバブルメディアを用いてオフラインでデータを移動させてもよい。

また、本実施の形態においては、電子メールから抽出された興味データにマッチする番組を推薦するようにしたが、これに限らず、例えば、ラジオ番組、またはインターネット上の Web サイト情報などを推薦することも勿論可能である。

さらに本発明によれば、番組を推薦する際に、エージェント 231 によって推薦理由を提示させるようにしたので、エージェント 231 に対して信頼性や親しみが感じられるようになる。

ところで、上述したエージェント 231 の表示および吹き出し 232 の中の台詞の表示、並びに、表示された台詞に対応する音声信号の出力については、本発明のエージェントプログラム 11 だけでなく、他のアプリケーション、例えば、ゲームやワードプロセッサのヘルプ画面等に適用することが可能である。さらに、ビデオカメラ、またはカーナビゲーション等のディスプレイに表示されるキャラクタに適用することも勿論可能である。

また、現在利用している機器とは、異なる機器において蓄積された嗜好データ 260 を用いて、現在利用している機器において番組の推薦を受けることも可能である。例えば、図 34 に示されるように、ネットワーク 5 に HDD レコーダ 2-1 および HDD レコーダ 2-2 に接続されるテレビジョン受像機 3-1、並びに HDD レコーダ 2-2 および HDD レコーダ 2-2 に接続されるテレビジョン受像機 3-2 が接続されている場合、ユーザは、HDD レコーダ 2-1 において、HDD レコーダ 2-2 に蓄積された嗜好データ 260 に基づいて、番組の推薦を受けることができる。

この場合の HDD レコーダ 2-1 における嗜好データ取得処理について、図 3 5 のフローチャートを参照して説明する。ステップ S 3 0 1 において、HDD レコーダ 2-1 は、データの取得が指示されたか否かを判定し、指示されたと判定されるまで待機する。指示の入力は、例えば、テレビジョン受像機 3 のモニタに
5 GUI が表示され、ユーザが入力部 7 6 を操作することにより行われる。

ステップ S 3 0 2 において、HDD レコーダ 2-1 は、取得機器の指定の入力を受け付ける。このとき、例えば、HDD レコーダ 2-2 を特定する機器 ID などが、ユーザにより入力される。ステップ S 3 0 3 において、HDD レコーダ 2-1 は、指定された機器の嗜好データを取得する。このとき、HDD レコーダ 2-1 は、ネットワーク 5 を介して HDD レコーダ 2-2 にアクセスし、HDD レコーダ 2-2 に蓄積された嗜好データ 2 6 0 を取得する。
10

ステップ S 3 0 4 において、HDD レコーダ 2-1 は、ステップ S 3 0 3 で取得された嗜好データ 2 6 0 をサーバ 6 に送信する。

このようにして、他の機器の嗜好データが取得され、サーバ 6 に送信される。

次に、ステップ S 3 0 4 において送信された嗜好データに基づく、サーバ 6 における番組情報検索処理について、図 3 6 のフローチャートを参照して説明する。ステップ S 3 2 1 において、サーバ 6 のイベント管理部 1 3 1 は、ネットワーク 5 を介して送信された、嗜好データ 2 6 0 を受信する。そして、イベント管理部 1 3 1 は、嗜好データ 2 6 0 をデータベース問い合わせ部 1 3 2 に供給する。
15

ステップ S 3 2 2 において、データベース問い合わせ部 1 3 2 は、嗜好データ 2 6 0 のフォーマットをチェックする。このとき、嗜好データの中に、番組情報を検索するために必要な情報（例えば、ジャンルなど）が含まれているかがチェックされる。
20

ステップ S 3 2 3 において、データベース問い合わせ部 1 3 2 は、データの補正が必要か否かを判定する。嗜好データの中に、番組情報を検索するために必要な情報が含まれている場合、データの補正は必要ないと判定され、ステップ S 3 2 5 に進む。一方、番組情報を検索するために必要な情報が含まれていない場合、
25

データの補正が必要と判定され、データベース問い合わせ部 1 3 2 は、ステップ S 3 2 4 に進み、図 3 8 のフローチャートを参照して後述するデータ補正処理を実行する。これにより、番組情報を検索するために必要な情報が追加され、嗜好データが補正される。

- 5 ステップ S 3 2 5 において、データベース問い合わせ部 1 3 2 は、推薦番組用データベースを検索し、嗜好データ 2 6 0 にマッチする番組情報を検索する。ステップ S 3 2 6 において、番組情報出力部 1 3 3 は、検索された番組情報を HDD レコーダ 2 - 1 に送信する。

10 このようにして、現在利用中の機器（HDD レコーダ 2 - 1）とは、異なる機器（HDD レコーダ 2 - 2）の嗜好データに基づいて、番組の推薦が行われる。このようにすることで、ユーザは、より興味の幅を広げるようにすることができる。

ところで、上述した嗜好データは、HDD レコーダ 2 - 1 または 2 - 2 において蓄積されるものに限らず、例えば、パーソナルコンピュータ 1 において蓄積された嗜好データとすることも可能である。図 3 7 は、パーソナルコンピュータ 1 において蓄積された嗜好データ 2 8 0 の構成例を示す図である。嗜好データ 2 8 0 は、嗜好データ 2 6 0 の構成と異なり、ジャンル 2 6 1、タイトル 2 6 2、放送時間帯 2 6 3、出演者情報 2 6 4 などの情報（図 3 3）が含まれず、キーワード 2 8 1 のみにより構成されている。

20 嗜好データ 2 8 0 には、番組情報を検索するために必要な情報（例えば、ジャンルなど）が含まれていないため、上述した図 3 6 のステップ S 3 2 3 において、データの補正が必要と判定され、ステップ S 3 2 4 のデータ補正処理が実行される。図 3 8 のフローチャートを参照して、図 3 6 のステップ S 3 2 4 のデータ補正処理について説明する。

25 ステップ S 3 4 1 において、データベース問い合わせ部 1 3 2 は、嗜好データ 2 8 0 からキーワードを抽出する。ステップ S 3 4 2 において、キーワードを辞書で検索する。辞書は、例えば、キーワードとジャンルが対応付けられたデータ

ベースであり、図 3 9 を参照して後述する辞書生成処理により、予めサーバ 6 の記憶部 5 9 に記憶されている。

ステップ S 3 4 3 において、データベース問い合わせ部 1 3 2 は、キーワードが一致したか否かを判定し、一致したと判定された場合、ステップ S 3 4 4 において、キーワードに対応するジャンルを取得する。ステップ S 3 4 5 において、データベース問い合わせ部 1 3 2 は、嗜好データ 2 8 0 にジャンルを追加し、データを補正する。

ステップ S 3 4 3 において、キーワードが一致していないと判定された場合、処理は、ステップ S 3 4 6 に進み、エラー情報が送信される。これにより HDD レコーダ 2 - 1 に対して、当該嗜好データに基づく検索ができない旨が通知される。

このようにすることで、番組情報を検索するために必要な情報が含まれていない嗜好データに基づいて、番組の推薦を行うことができる。特に、番組の視聴などを行わないパーソナルコンピュータ、携帯端末など機器における嗜好データに基づいて、番組の推薦を行うことが可能となり、よりユーザの興味の幅を広げることができる。

次に、図 3 9 のフローチャートを参照して、辞書作成処理について説明する。ステップ S 3 6 1 において、番組メタデータ取得部 1 2 1 は、メタデータを取得する。このとき、取得されるメタデータは、EPG データなどの番組情報でもよいし、ネットワーク 5 を介して取得されるコンテンツのメタデータであってもよい。そして、取得されたメタデータは、データ内容処理部 1 2 2 に転送される。図 4 0 は、データ内容処理部 1 2 2 の機能的構成例を示すブロック図である。この例では、メタデータの分析を行うメタデータ分析部 3 0 1 とメタデータ分析部 3 0 1 の分析結果に基づいて、辞書データの生成を行う辞書データ生成部 3 0 2 が設けられている。

ステップ S 3 6 2 において、メタデータ分析部 3 0 1 は、図 4 1 のフローチャートを参照して後述するメタデータ分析処理を実行する。これによりメタデータ

を構成する成分が抽出され、そのメタデータのジャンルと関連づけられて記憶される。ステップS 3 6 3において、辞書データ生成部3 0 2は、図4 4のフローチャートを参照して後述する辞書データ生成処理を実行する。これにより、キーワードとそのキーワードに対応するジャンルが記述された辞書データが生成される。

次に、図4 1のフローチャートを参照して、図4 0のステップS 3 6 2のメタデータ分析処理について説明する。ステップS 3 8 1において、メタデータ分析部3 0 1は、取得されたメタデータを成分に分解する。このとき、メタデータは、図4 2に示されるように分解される。

10 図4 2は、分解されたメタデータの例を示す図である。この例では、メタデータは、成分「ジャンル」、「放送局」、「放送時間帯」、「出演者」、「キーワード」に分解されている。成分「ジャンル」は、そのメタデータに対応するコンテンツのジャンルを表し、成分「放送局」は、そのメタデータに対応するコンテンツが放送される放送局を表し、成分「放送時間帯」は、そのメタデータに対応するコンテンツが放送される時間帯を表し、成分「出演者」は、そのメタデータ
15 に対応するコンテンツに登場する主な出演者を表し、成分「キーワード」は、そのメタデータに対応するコンテンツを紹介する文字情報などから所定の単語（例えば、名詞）が抽出されたものである。

第1番目のメタデータの成分「ジャンル」は、「料理」と記述され、成分「放送局」は「T A S」と記述され、成分「放送時間帯」は、「昼」と記述され、成分「出演者」は「A A A」と記述され、成分「キーワード」は、「レシピ、材料、作り方、文字放送、ステレオ、・・・」と記述されている。

第2番目のメタデータの成分「ジャンル」は、「生活情報」と記述され、成分「放送局」は「M H K」と記述され、成分「放送時間帯」は、「夜」と記述され、
25 成分「出演者」は「B B B」と記述され、成分「キーワード」は、「レジャー、行楽地、子供、文字放送、ステレオ、・・・」と記述されている。

第3番目のメタデータの成分「ジャンル」は、「子供」と記述され、成分「放

送局」は「MHK」と記述され、成分「放送時間帯」は、「朝」と記述され、成分「出演者」は「CCC」と記述され、成分「キーワード」は、「レジャー、子供、文字放送、ステレオ、・・・」と記述されている。

このようにして、メタデータが各成分に分解される。

- 5 ステップS 3 8 2において、メタデータ分析部 3 0 1は、メタデータのジャンルを検出する。ステップS 3 8 3において、メタデータ分析部 3 0 1は、ジャンルと各成分を対応付けて、RAM 7 3などの一時記憶装置に記憶する。このとき、メタデータは、ジャンル毎に集められ、図 4 3のように整理されて記憶される。図 4 3は、このとき、記憶されるデータの例を示す図である。この例では、成分
- 10 「ジャンル」に「料理」と記述されたメタデータが集められ、それぞれ上述した成分「ジャンル」、「放送局」、「放送時間帯」、「出演者」、「キーワード」に分解されて記述されている。同様に、「ジャンル」が「生活情報」と記述されたメタデータ、「ジャンル」が「子供」と記述されたメタデータなども、集められ記憶される。
- 15 このようにして、メタデータを構成する各成分が、メタデータのジャンル毎に集められ記憶される。

- 次に、図 4 4を参照して、図 3 9のステップS 3 6 3の辞書データ生成処理について説明する。ステップS 4 0 1において、辞書データ生成部 3 0 2は、ステップS 3 8 3で記憶されたメタデータの中で、各放送局の各時間帯に含まれるキ
- 20 ーワードを検出する。例えば、図 4 2において成分「キーワード」に含まれる「文字放送」、「ステレオ」は、各放送局の各時間帯において検出される。このような単語（キーワード）は、コンテンツの内容を把握する上で重要ではない単語、すなわちノイズと考えられる。そこで、ステップS 4 0 2において、辞書データ生成部 3 0 2は、該当するキーワードをノイズとして削除する。
- 25 ステップS 4 0 3において、辞書データ生成部 3 0 2は、ジャンル毎に共起率の高いキーワードを検出する。例えば、図 4 3の成分「キーワード」に含まれる「レシピ」、「材料」、「作り方」は、「ジャンル」が「料理」に分類された、

第1番目乃至第3番目のメタデータに共通して含まれている。このような単語（キーワード）は、「ジャンル」が「料理」に分類されるメタデータにおいて共起率の高いキーワードとして検出される。

5 ステップS404において、辞書データ生成部302は、ジャンルとキーワードを関連付けて、辞書データとして記憶する。図45に、このとき記憶される辞書データの例を示す。この例では、辞書データは、「キーワード」、「頻度／月」、「ジャンル」、「他の成分」により構成されている。「キーワード」には、ステップS403で検出された共起率のキーワードが記述される。この例では、
10 「レシピ」、「材料」、「作り方」、・・・が記述されている。「頻度／月」には、そのキーワードが1ヶ月間に検出された回数が記述される。「頻度／月」に記述された数値が大きいキーワードは、現在流行しているキーワードと考えられる。

 「ジャンル」には、そのキーワードが属するジャンルが記述される。例えば第1番目のキーワード「レシピ」、第2番目のキーワード「材料」、第3番目のキーワード「作り方」は、どれも「ジャンル」が「料理」と記述されている。また、
15 第5番目のキーワード「行楽地」は、「ジャンル」が生活情報と記述されている。第4番目のキーワード「レジャー」と第6番目のキーワード「子供」において「ジャンル」として「生活情報」と「子供」が記述されている。これは、キーワード「レジャー」（または「子供」）は、そのジャンルが「生活情報」と分類されるメタデータにおいて共起率が高いと同時に、そのジャンルが「子供」と分類
20 されるメタデータにおいても共起率が高いことを表す。

 「他の成分」には、そのキーワードと同様に、そのジャンルにおいて共起率が高いと判定された成分が記述される。

 このようにして、辞書データが生成され、データベース化され、辞書として記憶される。このようにして作成された辞書を使用することにより、上述したように、番組情報を検索するために必要な情報が含まれていない嗜好データに基づいて、番組の推薦を行うことができる。また、番組情報などのメタデータのデータ
25

ベースを作成するとき、辞書を参照することにより、ジャンルが付与されていないメタデータにジャンルを付与して番組情報を生成したり、現在流行している番組の番組情報を生成することができる。

- 次に、図 4 6 を参照して、辞書を使用してメタデータのデータベースを作成するデータベース作成処理について説明する。ステップ S 4 2 1 において、番組メタデータ取得部 1 2 1 は、転送されたメタデータを取得する。そして、取得されたメタデータは、データ内容処理部 1 2 2 に転送される。

- ステップ S 4 2 2 において、データ内容処理部 1 2 2 は、取得されたメタデータにジャンルが付与されているか否かを判定し、ジャンルが付与されていないと判定された場合、ステップ S 4 2 3 に進み、メタデータのキーワードを検出する。ステップ S 4 2 4 において、データ内容処理部 1 2 2 は、キーワードがあった（取得できた）か否かを判定し、キーワードがあったと判定された場合、ステップ S 4 2 5 に進み、取得されたキーワードで辞書データを検索する。

- ステップ S 4 2 6 において、データ内容処理部 1 2 2 は、辞書データの中に、一致するキーワードがあったか否かを判定し、一致するキーワードがなかったと判定された場合、ステップ S 4 2 8 に進み、別のキーワードを検出し、ステップ S 4 2 4 に戻る。

ステップ S 4 2 4 において、キーワードがなかった（取得できなかった）と判定された場合、処理は終了される。

- ステップ S 4 2 6 において、一致するキーワードがあったと判定された場合、ステップ S 4 2 7 に進み、データ内容処理部 1 2 2 は、キーワードに対応するジャンルを取得する。例えば、ステップ S 4 2 3 またはステップ S 4 2 8 の処理でキーワード「レシピ」が検出された場合、図 4 5 を参照して上述したように、このキーワードに対応するジャンルは「料理」と判定され、このメタデータに対応するジャンルとして「料理」が取得される。

ステップ S 4 2 9 において、データ内容処理部 1 2 2 は、図 4 7 を参照して後述するデータ記述処理を実行する。これにより、番組情報（メタデータ）のデー

データベースが記憶される。

一方、ステップ S 4 2 2 において、当該メタデータにジャンルが付与されていると判定された場合、ステップ S 4 3 0 に進み、データ内容処理部 1 2 2 は、そのジャンルを取得し、ステップ S 4 2 9 に進み、データ記述処理を実行する。

- 5 このようにして、取得されたメタデータに基づいて、番組情報（メタデータ）のデータベースが作成される。ジャンルが付与されていないメタデータについても、そのジャンルが付与され、データベースに記憶される。

次に、図 4 7 を参照して、図 4 6 のステップ S 4 2 9 のデータ記述処理について説明する。ステップ S 4 5 1 において、データ内容処理部 1 2 2 は、メタデー

- 10 タの成分を補完する。

例えば、メタデータの成分が複数ある場合、特定の成分間の相関が極端に高く、他の成分との相関が極端に低い組み合わせを抽出することにより、その組み合わせを用いて、部分的に欠如している成分を補完することができる。例えば、メタデータの成分として A, B, C, D...X があり、それぞれの属性値が成分 A については、A1, A2, A3, 成分 B について、B1, B2, B3, B4, 成分 C について、C1, C2, 成分 D について、D1, D2, D3...のように構成されていたとする。

- 15 取得済のメタデータについては、図 4 5 の辞書データの「他の成分」を参照し、各成分の相関を調べることができる。これにより、成分 A1 と B3、および成分 C2 と D2 にだけ強い相関があり、他には相関がないことが判明したとする。このとき、ある新しいコンテンツのメタデータを取得し、そのメタデータの成分 A と D が付与されていない状態であり、かつ成分 B は B3 であり、成分 C は C2 であった場合、そのメタデータを構成する成分は、A1, B3, C2, D2 と高い確立で予測することができる。このようにして、未付与だった成分 A と D を付与することができ、メタデータの成分が補完される。

- 25 ステップ S 4 5 2 において、データ内容処理部 1 2 2 は、ステップ S 4 2 3 または S 4 2 8 で検出された当該キーワードは、頻度大か否かを判定する。このとき、辞書データにおいて、当該キーワードに対応する「頻度／月」の数値が検出

され、その数値が所定の値（例えば、10）以上である場合、当該キーワードは頻度大であると判定される。

ステップS452において、当該キーワードは頻度大であると判定された場合、ステップS453に進み、当該メタデータの流行区分を流行と設定する。このように、メタデータに流行区分を設けて関連情報として記憶しておくことにより、その関連情報を用いて、ユーザに対して、流行のコンテンツを推薦することができる。

一方、ステップS452において、当該キーワードは頻度大ではないと判定された場合、ステップS453の処理は、スキップされる。

10 ステップS454において、データ内容処理部122は、当該メタデータとジャンルを関連付けて、データベースとして記憶する。

このようにして、メタデータを構成する成分が補完され、また流行区分が設定されて、データベースに記述される。

15 上述した一連の処理を実行するエージェントプログラム11またはサーバプログラム101は、パーソナルコンピュータに予め組み込まれるか、あるいは、記録媒体からインストールされる。

20 上述した一連の処理は、ハードウェアに実行させることもできるが、通常、ソフトウェアにより実行させる。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するエージェントプログラム1が、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、記録媒体からインストールされる。

25 コンピュータにインストールされ、コンピュータによって実行可能な状態とされるプログラムを記録する記録媒体は、図3または図6に示されるように、プログラムが記録されている磁気ディスク62、150（フレキシブルディスクを含む）、光ディスク63、151（CD-ROM（Compact Disk-Read Only Memory）、DVD（Digital Versatile Disk）を含む）、光磁気ディスク64、152

(MD (Mini-Disk)を含む)、もしくは半導体メモリ 65, 153 などよりなるパッケージメディア、または、プログラムが一時的もしくは永続的に記録される ROM 52, 142 や記憶部 59, 147 を構成するハードディスクなどにより構成される。記録媒体に対するプログラムの記録は、必要に応じてルータ、モデム
5 などのインタフェースを介して、公衆回線網、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の通信媒体を利用して行われる。

なお、本明細書において、記録媒体に記録されるプログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時
10 系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。

また、本明細書において、システムとは、複数の装置により構成される装置全体を表わすものである。

15 産業上の利用可能性

以上のように、第1の本発明によれば、テレビ番組の推薦を行うことができる。

また第1の本発明によれば、ユーザがやりとりした電子メールから興味を抽出し、その興味にマッチしたテレビ番組の推薦を行うことができる。

第2の本発明によれば、容易にテレビ番組の検索を要求することができる。

20 また第2の本発明によれば、ユーザがやりとりした電子メールから興味を抽出し、その興味にマッチしたテレビ番組の検索を要求することができる。

第3の本発明によれば、データベース化された番組情報から、テレビ番組を検索することができる。

25 また第3の本発明によれば、ユーザがやりとりした電子メールから抽出された興味に基づいて、その興味にマッチしたテレビ番組をデータベースから検索して推薦することができる。

第4の本発明によれば、外出先においても簡単にテレビ番組の録画予約を行うことができる。

また第4の本発明によれば、外出先においてテレビ番組の録画予約を行うための電子メールを作成し、自宅にある録画装置に送信する。そして、録画装置は、

5 テレビ番組の録画予約を行うための電子メールを受信し、その電子メールから予約情報を抽出し、予約情報にマッチしたテレビ番組の番組情報をサーバから取得して録画予約を行うことができる。

第5の本発明によれば、簡単にテレビ番組の録画予約を行うことができる。

また第5の本発明によれば、テレビ番組の録画予約を行うための電子メールを

10 受信し、その電子メールから予約情報を抽出し、予約情報にマッチしたテレビ番組の番組情報をサーバから取得して録画予約を行うことができる。

第6の本発明によれば、ユーザがやりとりした電子メールから興味を抽出し、その興味にマッチしたテレビ番組の推薦を行うことができる。

第7の本発明によれば、容易にユーザがやりとりした電子メールから興味を抽

15 出し、その興味にマッチしたテレビ番組の検索を要求することができる。

請求の範囲

1. 情報処理装置、およびネットワークを介して前記情報処理装置からアクセスされる情報検索装置からなる情報検索システムにおいて、

前記情報処理装置は、

- 5 所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

前記抽出手段により抽出された前記興味語を前記情報検索装置に送信し、

前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段に基づいて、前記情報検索装置から送信されてくる前記番

- 10 組情報を受信する受信手段と

を備え、

前記情報検索装置は、

前記番組情報を蓄積する蓄積手段と、

前記情報処理装置から送信されてきた前記検索要求に基づいて、前記検索要

- 15 求に含まれる前記興味語に関連する前記番組情報を前記蓄積手段から検索する検索手段と、

前記検索手段により検索された前記番組情報を前記情報処理装置に送信する送信手段と

を備えることを特徴とする情報検索システム。

- 20 2. 前記情報処理装置の前記抽出手段は、前記所定の情報を形態素解析して前記興味語に分解する形態素解析手段を含む

ことを特徴とする請求の範囲第1項に記載の情報検索システム。

3. 前記情報処理装置は、前記抽出手段により抽出された前記興味語を用いてデータベースを作成するデータベース構築手段をさらに備える

- 25 ことを特徴とする請求の範囲第1項に記載の情報検索システム。

4. 前記情報処理装置は、前記受信手段により受信された前記番組情報に基づいて、前記番組の録画を制御する録画制御手段をさらに備える

ことを特徴とする請求の範囲第1項に記載の情報検索システム。

5. 前記情報処理装置は、前記受信手段により受信された前記番組情報の表示を制御する表示制御手段をさらに備える

ことを特徴とする請求の範囲第1項に記載の情報検索システム。

5 6. 前記情報検索装置の前記蓄積手段は、前記番組情報を用いて、前記番組に関連付けてデータベースを作成するデータベース構築手段を含む

ことを特徴とする請求の範囲第1項に記載の情報検索システム。

7. 前記所定の情報は、文書情報、前記番組に関する嗜好情報、または、前記番組の視聴履歴に関する情報のうち、少なくとも1つを含む

10 ことを特徴とする請求の範囲第1項に記載の情報検索システム。

8. 前記文書情報は、電子メールである

ことを特徴とする請求の範囲第7項に記載の情報検索システム。

9. 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含む

15 ことを特徴とする請求の範囲第1項に記載の情報検索システム。

10. 前記情報処理装置は、他の情報処理装置から前記所定の情報を取得する

ことを特徴とする請求の範囲第1項に記載の情報検索システム。

11. 所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

20 前記抽出手段により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段に基づいて、前記情報検索装置から送信されてくる前記番組情報を受信する受信手段と

を備えることを特徴とする情報処理装置。

25 12. 前記抽出手段は、前記所定の情報を形態素解析して前記興味語に分解する形態素解析手段を含む

ことを特徴とする請求の範囲第11項に記載の情報処理装置。

1 3. 前記抽出手段により抽出された前記興味語を用いてデータベースを作成するデータベース構築手段をさらに備える

ことを特徴とする請求の範囲第 1 1 項に記載の情報処理装置。

1 4. 前記受信手段により受信された前記番組情報に基づいて、前記番組の録
5 画を制御する録画制御手段をさらに備える

ことを特徴とする請求の範囲第 1 1 項に記載の情報処理装置。

1 5. 前記受信手段により受信された前記番組情報の表示を制御する表示制御手段をさらに備える

ことを特徴とする請求の範囲第 1 1 項に記載の情報処理装置。

10 1 6. 前記所定の情報は、文書情報、前記番組に関する嗜好情報、または、前記番組の視聴履歴に関する情報のうち、少なくとも 1 つを含む

ことを特徴とする請求の範囲第 1 1 項に記載の情報処理装置。

1 7. 前記文書情報は、電子メールである

ことを特徴とする請求の範囲第 1 6 項に記載の情報処理装置。

15 1 8. 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含む

ことを特徴とする請求の範囲第 1 1 項に記載の情報処理装置。

1 9. 前記所定の情報は、他の情報処理装置から取得される

ことを特徴とする請求の範囲第 1 1 項に記載の情報処理装置。

20 2 0. 番組の録画を制御する録画制御手段を備える情報処理装置の情報処理方法において、

前記所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、

25 前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

を含むことを特徴とする情報処理方法。

21. 番組の録画を制御する録画制御手段を備える情報処理装置を制御するプログラムであって、

5 前記所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

10 を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

22. 番組の録画を制御する録画制御手段を備える情報処理装置を制御するコンピュータに、

15 前記所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

20 を実行させるプログラム。

23. 番組に関する番組情報を蓄積する蓄積手段と、

情報処理装置から送信されてきた、前記番組情報を取得するための興味語を受信する受信手段と、

25 前記受信手段により受信された前記興味語に関連する前記番組情報を前記蓄積手段から検索する検索手段と、

前記検索手段により検索された前記番組情報を前記情報処理装置に送信する送信手段と

を備えることを特徴とする情報検索装置。

24. 前記興味語は、前記情報処理装置において所定の情報が形態素解析されて分解された単語である

ことを特徴とする請求の範囲第23項に記載の情報検索装置。

5 25. 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含む

ことを特徴とする請求の範囲第23項に記載の情報検索装置。

26. 前記番組情報を分析する分析手段と、

前記分析手段の分析結果に基づいて、前記番組情報のジャンルとキーワードを

10 対応付ける辞書データを生成する辞書作成手段と、

前記辞書作成手段により生成された辞書データに基づいて、前記番組情報にジャンルを付与して記憶するデータベース作成手段と

をさらに備えることを特徴とする請求の範囲第23項に記載の情報検索装置。

27. 前記興味語からキーワードを抽出し、前記キーワードに基づいて、前記

15 辞書データを検索して前記キーワードに対応するジャンルを取得し、前記ジャンルに基づいて、前記番組情報を検索するキーワード検索手段をさらに備える

ことを特徴とする請求の範囲第23項に記載の情報検索装置。

28. 前記辞書作成手段は、

前記メタデータに含まれる単語のうち、特定のジャンルのメタデータにおいて

20 共起率の高い単語を、そのジャンルのキーワードとして検出するキーワード検出手段を備える

ことを特徴とする請求の範囲第26項に記載の情報検索装置。

29. 前記辞書作成手段は、前記キーワードが検出される頻度を前記キーワードとともに記憶し、前記辞書データを生成する

25 ことを特徴とする請求の範囲第26項に記載の情報検索装置。

30. 前記データベース作成手段は、前記番組情報に含まれる成分に基づいて、前記番組情報に含まれていない成分を補完する

ことを特徴とする請求項の範囲第 26 項に記載の情報検索装置。

31. 情報を検索する情報検索装置の情報検索方法において、

番組に関する番組情報の蓄積を制御する蓄積制御ステップと、

情報処理装置から送信されてきた、前記番組情報を取得するための興味語の受

5 信を制御する受信制御ステップと、

前記受信制御ステップの処理により受信が制御された前記興味語に関連する前記番組情報を検索する検索ステップと、

前記検索ステップの処理により検索された前記番組情報の、前記情報処理装置への送信を制御する送信制御ステップと

10 を含むことを特徴とする情報検索方法。

32. 情報を検索する情報検索装置を制御するプログラムであって、

番組に関する番組情報の蓄積を制御する蓄積制御ステップと、

情報処理装置から送信されてきた、前記番組情報を取得するための興味語の受信を制御する受信制御ステップと、

15 前記受信制御ステップの処理により受信が制御された前記興味語に関連する前記番組情報を検索する検索ステップと、

前記検索ステップの処理により検索された前記番組情報の、前記情報処理装置への送信を制御する送信制御ステップと

20 を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

33. 情報を検索する情報検索装置を制御するコンピュータに、

番組に関する番組情報の蓄積を制御する蓄積制御ステップと、

情報処理装置から送信されてきた、前記番組情報を取得するための興味語の受信を制御する受信制御ステップと、

25 前記受信制御ステップの処理により受信が制御された前記興味語に関連する前記番組情報を検索する検索ステップと、

前記検索ステップの処理により検索された前記番組情報の、前記情報処理装置への送信を制御する送信制御ステップと

を実行させるプログラム。

- 3 4. 携帯端末装置、前記携帯端末装置にネットワークを介して接続される情報処理装置、および、前記ネットワークを介して前記情報処理装置からアクセスされる情報検索装置からなる情報検索システムにおいて、

前記携帯端末装置は、

番組を録画予約するための予約情報を生成する生成手段と、

前記生成手段により生成された前記予約情報を前記情報処理装置に送信する

- 10 第1の送信手段と

を備え、

前記情報処理装置は、

前記携帯端末装置から送信されてきた前記予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

- 15 前記抽出手段により抽出された前記興味語を前記情報検索装置に送信し、
前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段に基づいて、前記情報検索装置から送信されてくる前記番組情報を受信する受信手段と

を備え、

- 20 前記情報検索装置は、

前記番組情報を蓄積する蓄積手段と、

前記情報処理装置から送信されてきた前記検索要求に基づいて、前記検索要求に含まれる前記興味語に関連する前記番組情報を前記蓄積手段から検索する検索手段と、

- 25 前記検索手段により検索された前記番組情報を前記情報処理装置に送信する
第2の送信手段と

を備えることを特徴とする情報検索システム。

35. 前記情報処理装置の前記抽出手段は、前記所定の情報を形態素解析して前記興味語に分解する形態素解析手段を含む

ことを特徴とする請求の範囲第34項に記載の情報検索システム。

36. 前記情報処理装置は、前記受信手段により受信された前記番組情報に基づいて、前記番組の録画を制御する録画制御手段をさらに備える

ことを特徴とする請求の範囲第34項に記載の情報検索システム。

37. 前記情報検索装置の前記蓄積手段は、前記番組情報を用いて、前記番組に関連付けてデータベースを作成するデータベース構築手段を含む

ことを特徴とする請求の範囲第34項に記載の情報検索システム。

10 38. 前記予約情報は、番組名、ジャンル名、または、出演者名のうち、少なくとも1つを含む

ことを特徴とする請求の範囲第34項に記載の情報検索システム。

39. 前記予約情報は、電子メールである

ことを特徴とする請求の範囲第34項に記載の情報検索システム。

15 40. 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含む

ことを特徴とする請求の範囲第34項に記載の情報検索システム。

41. 携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出手

20 段と、

前記抽出手段により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段に基づいて、前記情報検索装置から送信されてくる前記番組情報を受信する受信手段と

25 を備えることを特徴とする情報処理装置。

42. 前記抽出手段は、前記所定の情報を形態素解析して前記興味語に分解する形態素解析手段を含む

ことを特徴とする請求の範囲第41項に記載の情報処理装置。

43. 前記受信手段により受信された前記番組情報に基づいて、前記番組の録画を制御する録画制御手段をさらに備える

ことを特徴とする請求の範囲第41項に記載の情報処理装置。

5 44. 前記予約情報は、番組名、ジャンル名、または、出演者名のうち、少なくとも1つを含む

ことを特徴とする請求の範囲第41項に記載の情報処理装置。

45. 前記予約情報は、電子メールである

ことを特徴とする請求の範囲第41項に記載の情報処理装置。

10 46. 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含む

ことを特徴とする請求の範囲第41項に記載の情報処理装置。

47. 番組の録画を制御する録画制御手段を備える情報処理装置の情報処理方法において、

15 携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

20 前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

を含むことを特徴とする情報処理方法。

48. 番組の録画を制御する録画制御手段を備える情報処理装置を制御するプログラムであって、

25 携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、
前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる
前記番組情報の受信を制御する受信制御ステップと

5 を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

49. 番組の録画を制御する録画制御手段を備える情報処理装置を制御するコンピュータに、

10 携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、
前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

15 前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと
 を実行させるプログラム。

50. 情報処理装置、およびネットワークを介して前記情報処理装置からアクセスされる情報検索装置からなる情報検索システムにおいて、

前記情報処理装置は、

20 電子メールを分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

前記抽出手段により抽出された前記興味語を前記情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

25 前記検索要求手段による検索要求に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信する受信手段と
 を備え、

前記情報検索装置は、

前記番組情報を蓄積する蓄積手段と、

前記情報処理装置から送信されてきた前記検索要求に基づいて、前記検索要求に含まれる前記興味語に関連する前記番組情報を前記蓄積手段から検索する検索手段と、

- 5 前記検索手段により検索された前記番組情報を前記情報処理装置に送信する送信手段と

を備えることを特徴とする情報検索システム。

5 1. 電子メールを分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

- 10 前記抽出手段により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段による検索要求に基づいて、前記情報検索装置から送信されてくる前記番組情報を受信する受信手段と

を備えることを特徴とする情報処理装置。

- 15 5 2. 番組の録画を制御する録画制御手段を備える情報処理装置の情報処理方法において、

電子メールを分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、

- 20 前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理による検索要求に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

を含むことを特徴とする情報処理方法。

- 25 5 3. 番組の録画を制御する録画制御手段を備える情報処理装置を制御するプログラムであって、

電子メールを分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理による検索要求に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

5 ．を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

5 4 ． 番組の録画を制御する録画制御手段を備える情報処理装置を制御するコンピュータに、

電子メールの情報を分析し、番組に関する番組情報を取得するための興味語を

10 抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理による検索要求に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

15 ．を実行させるプログラム。

1/41

図1

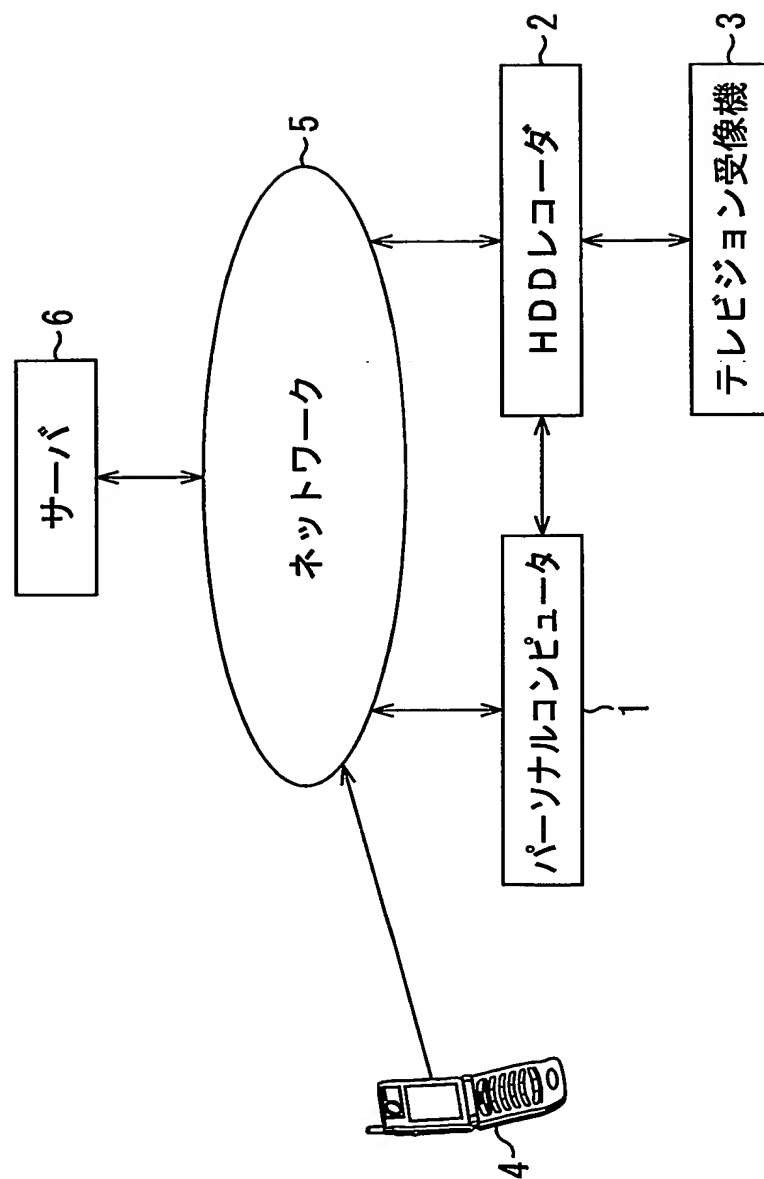
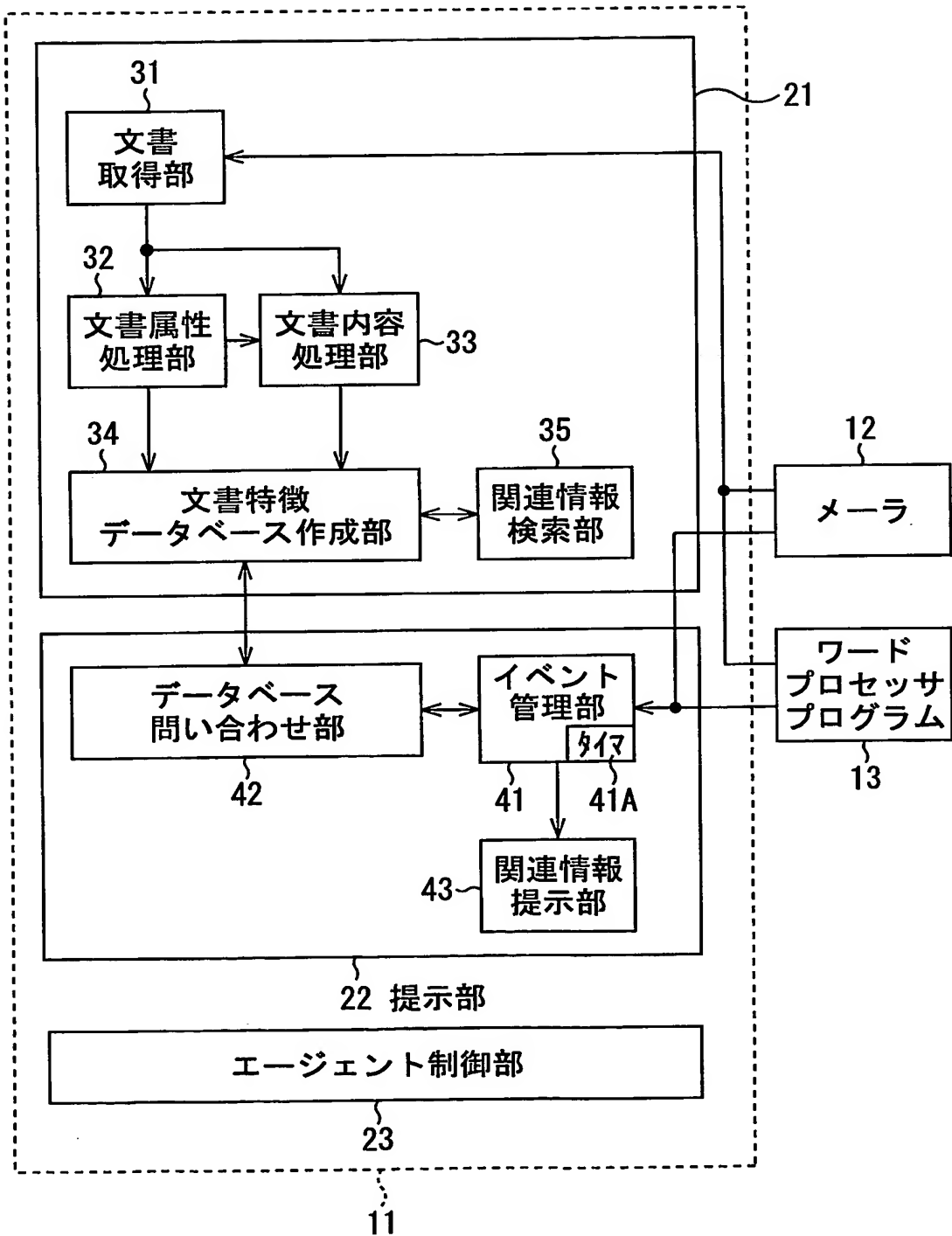


図 2



3/41

図3

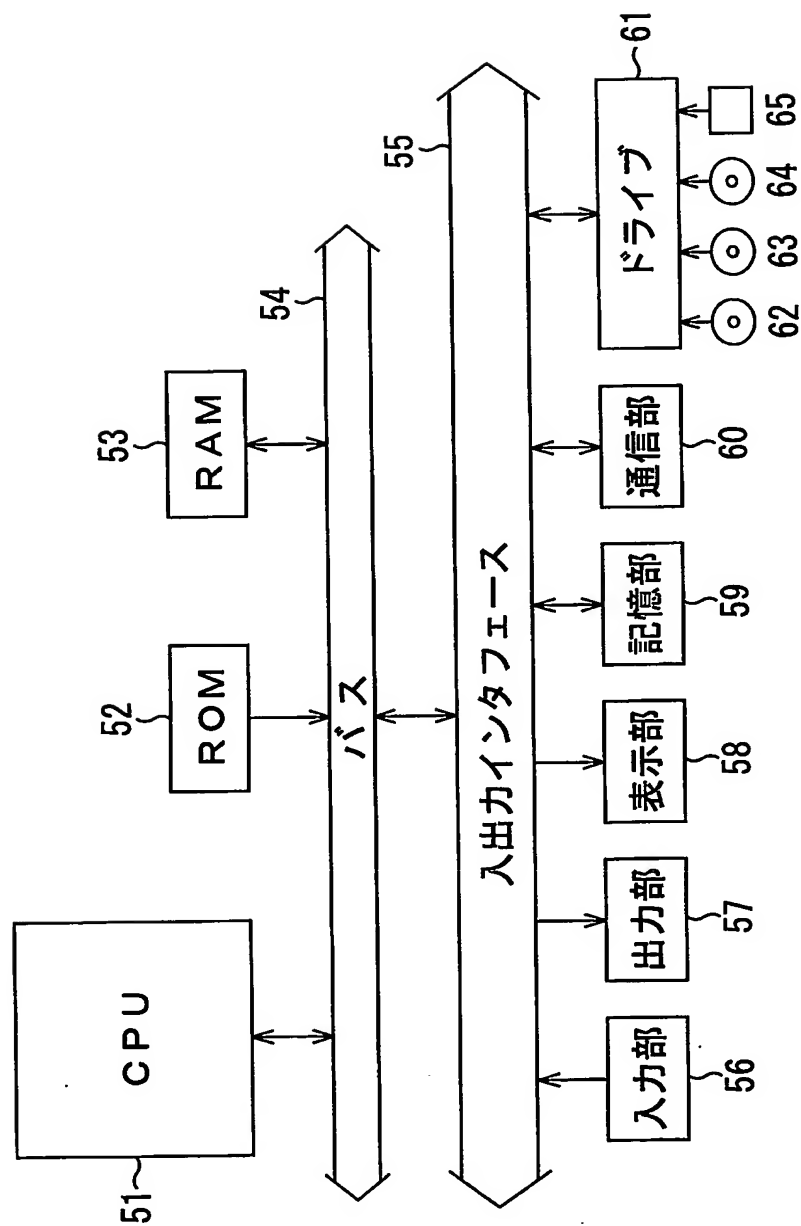
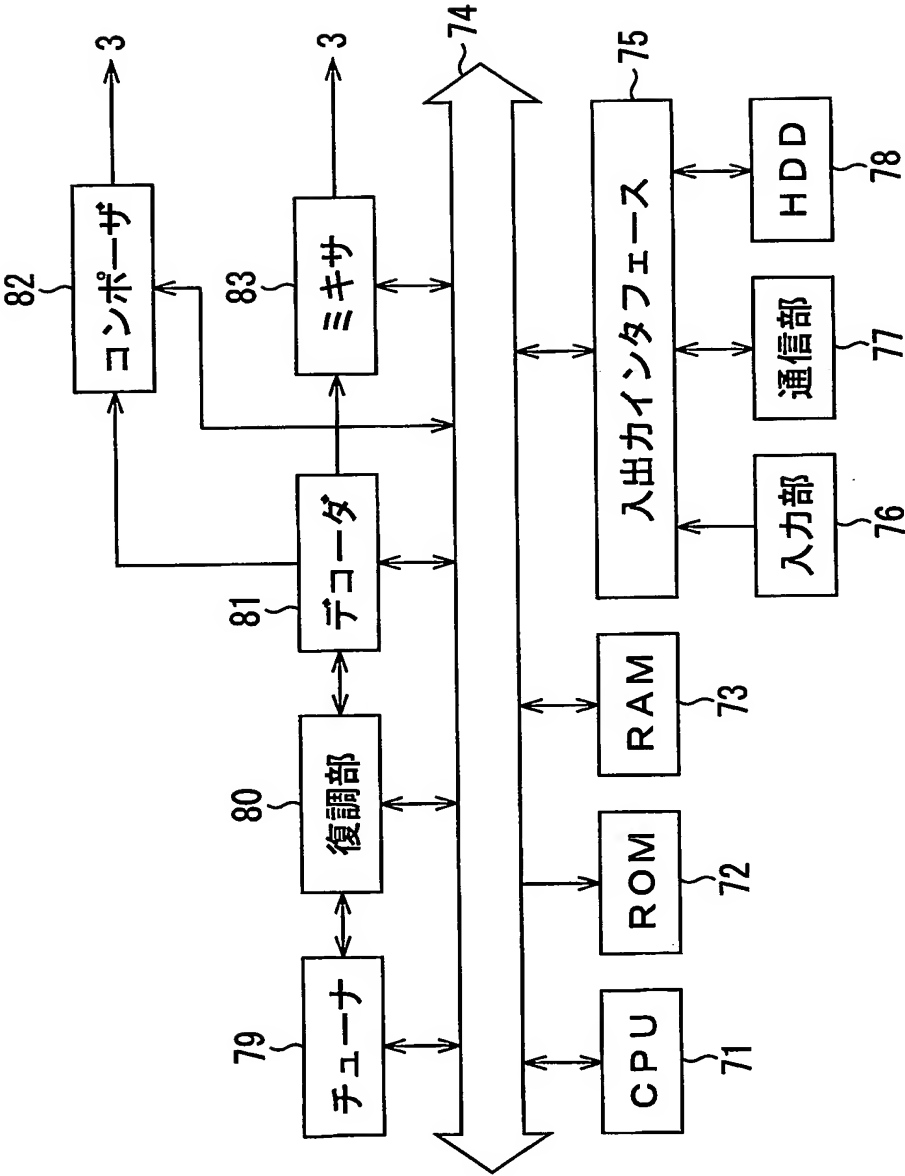
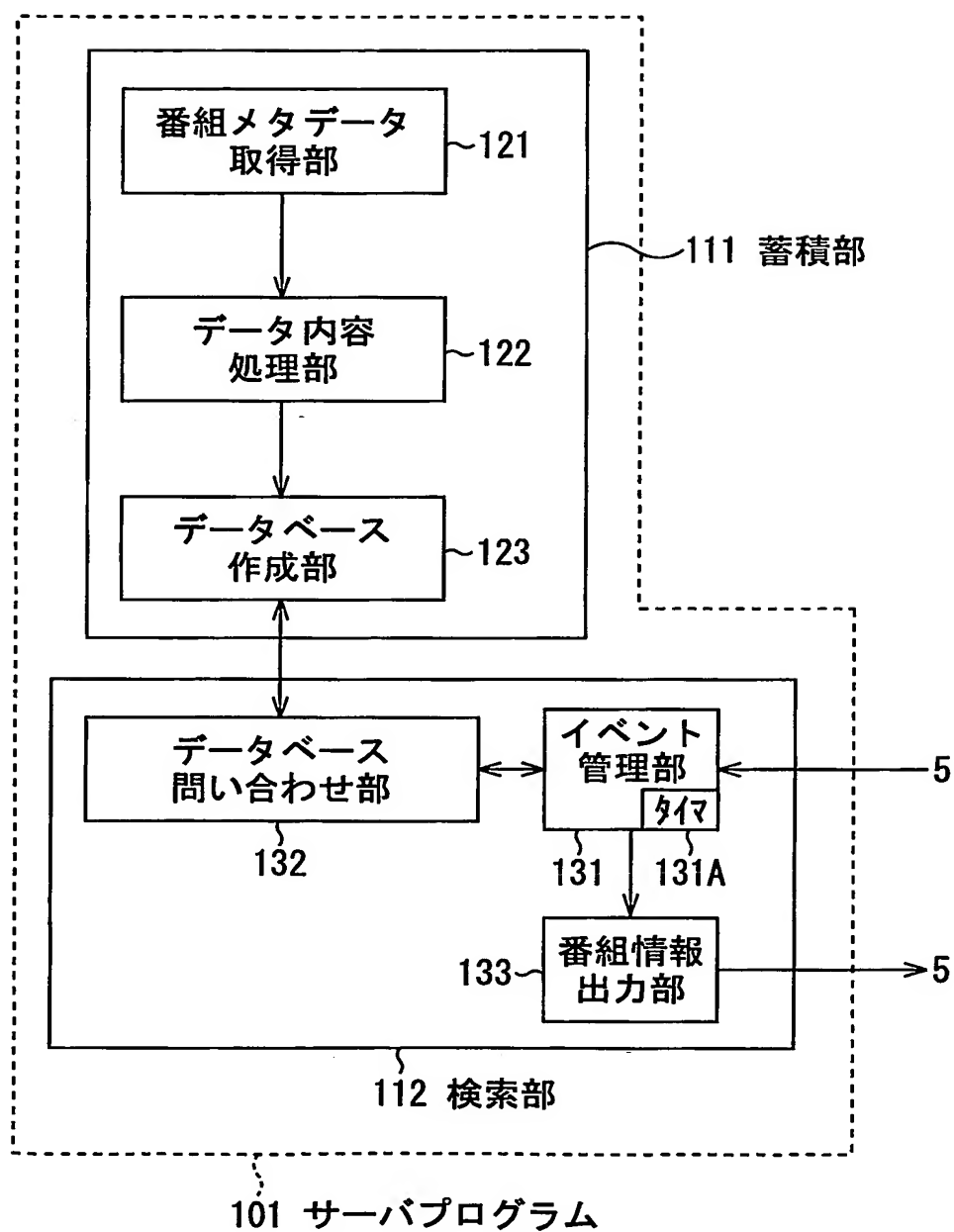


図4



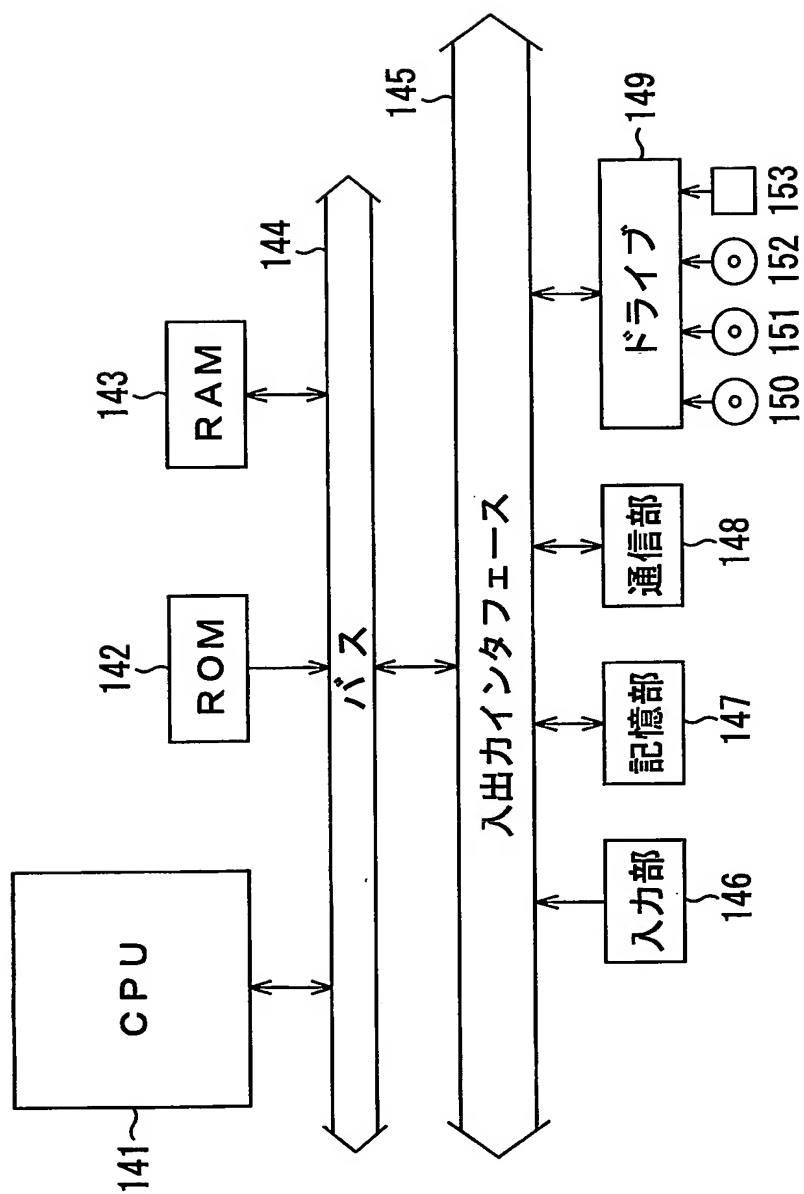
5/41

図 5



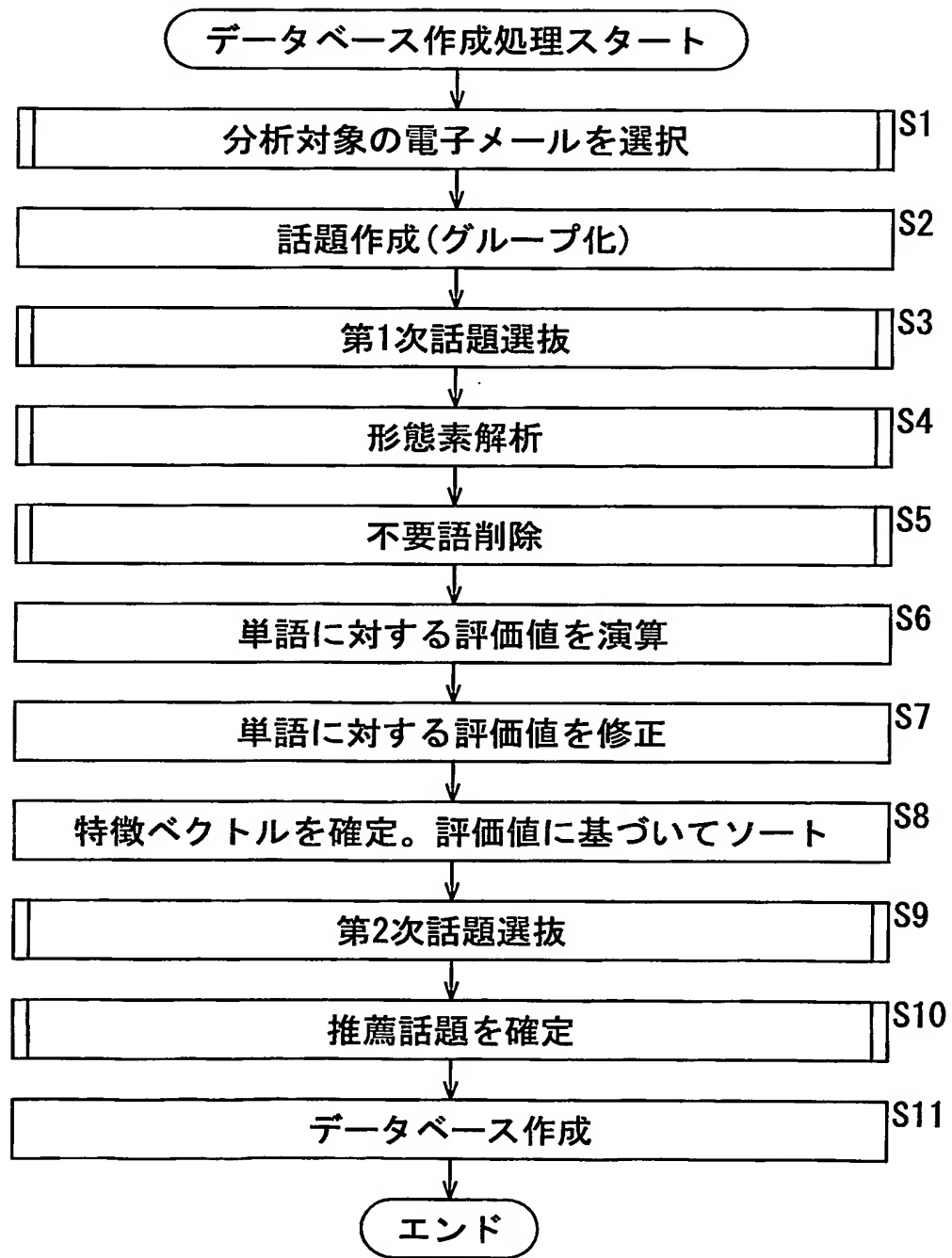
6/41

図 6



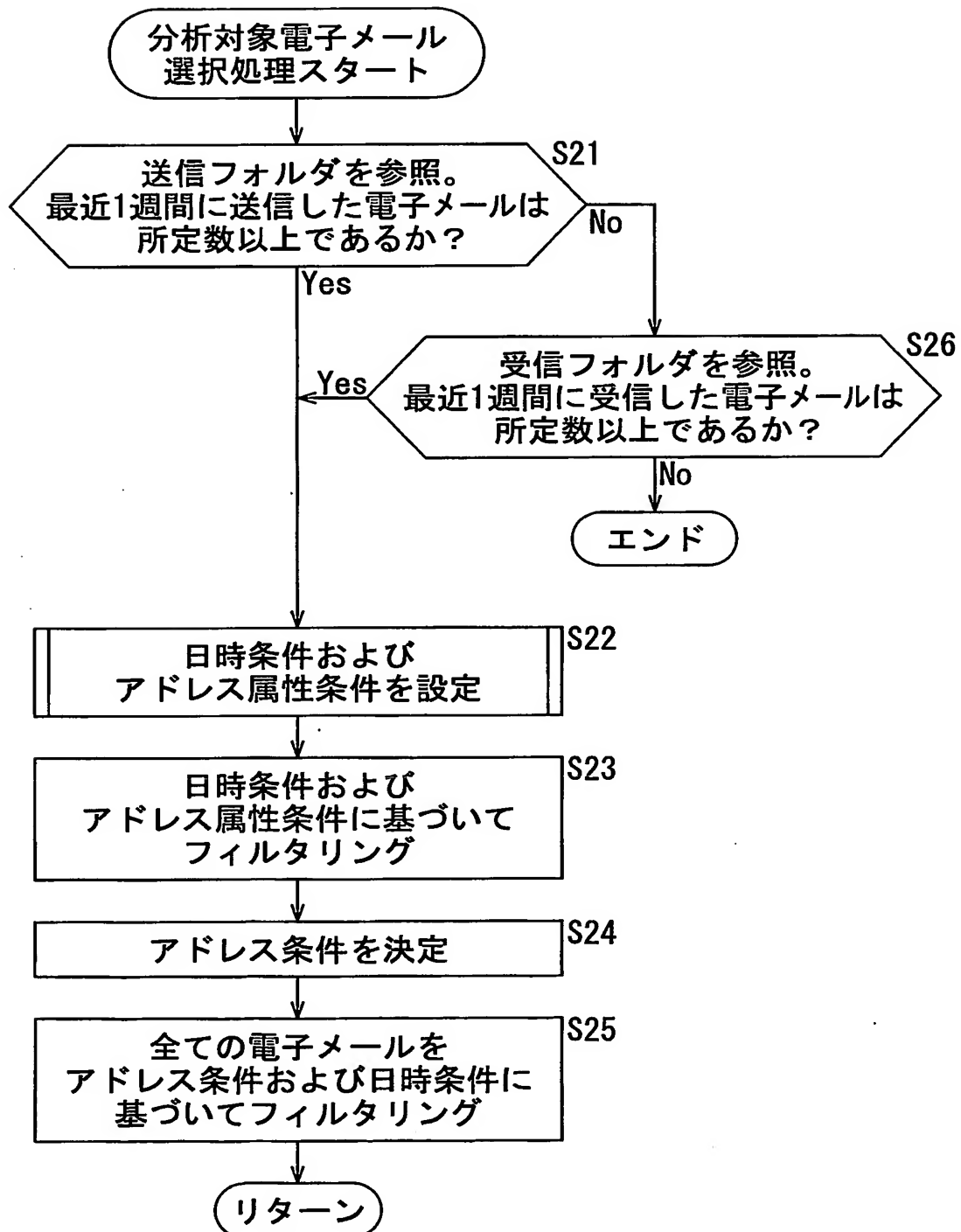
7/41

図 7



8/41

図 8



9/41

図 9

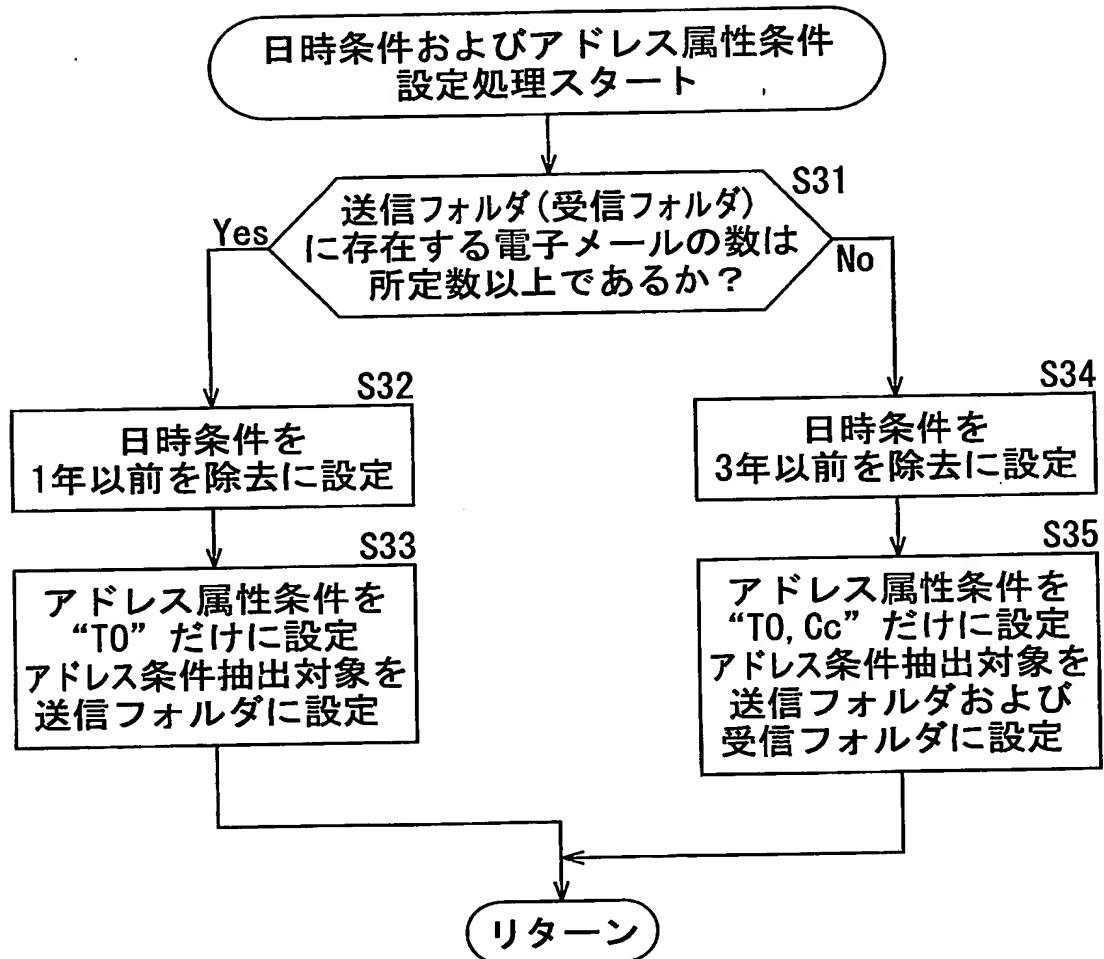


図10

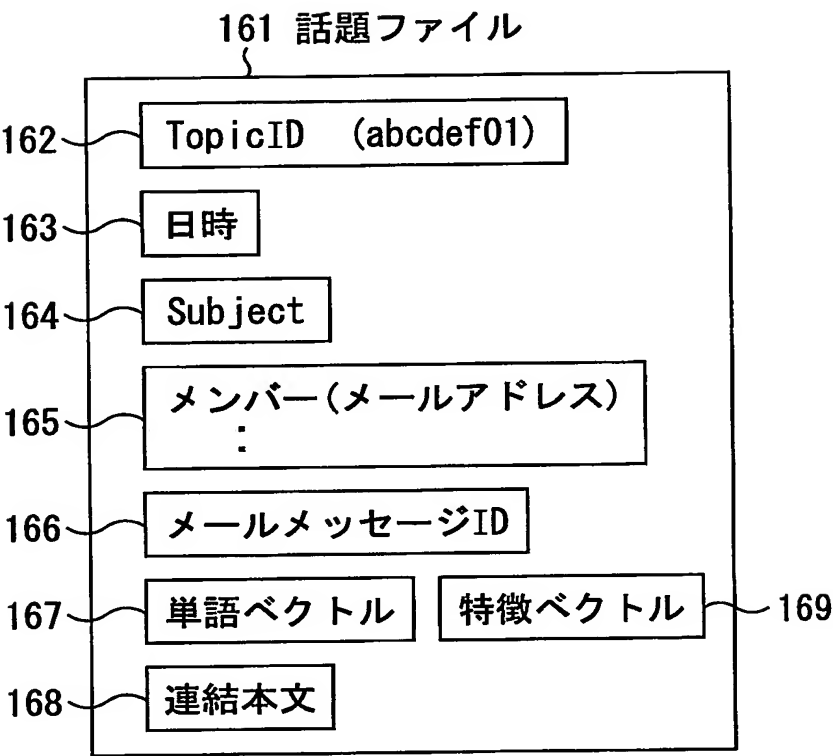
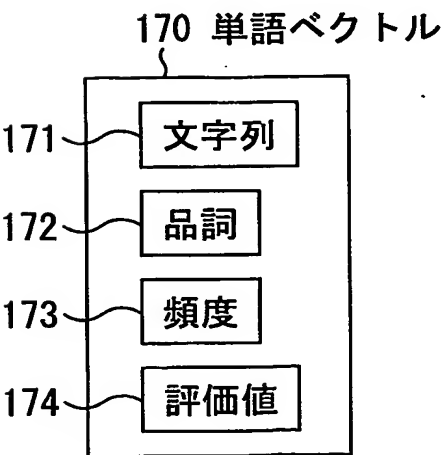
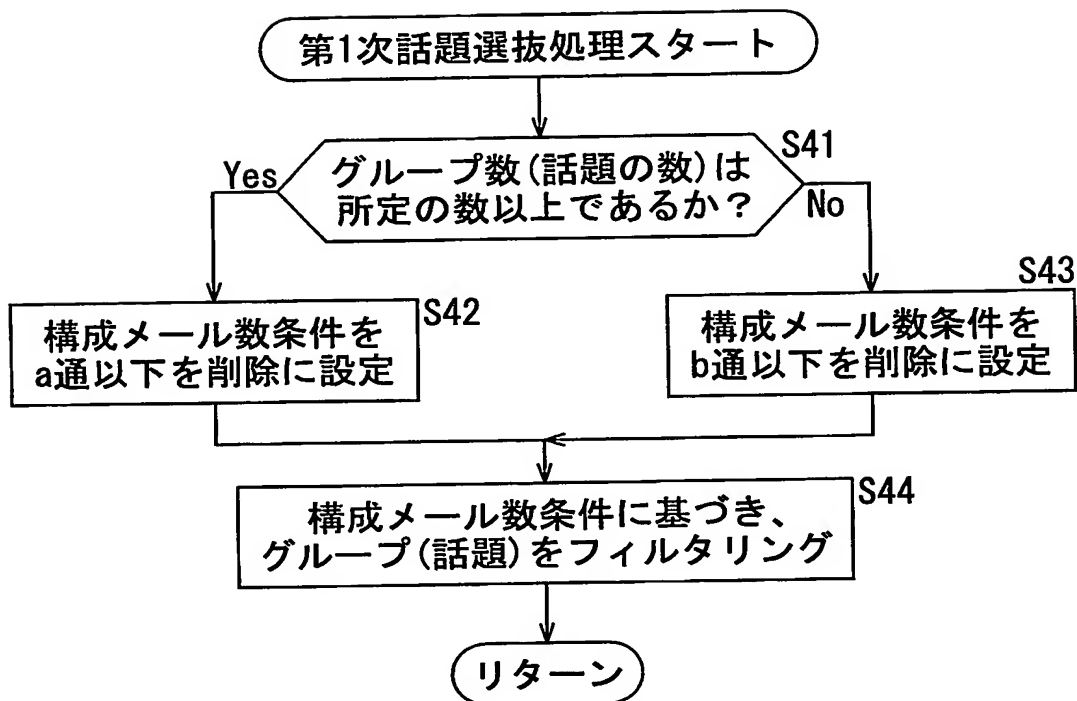


図11



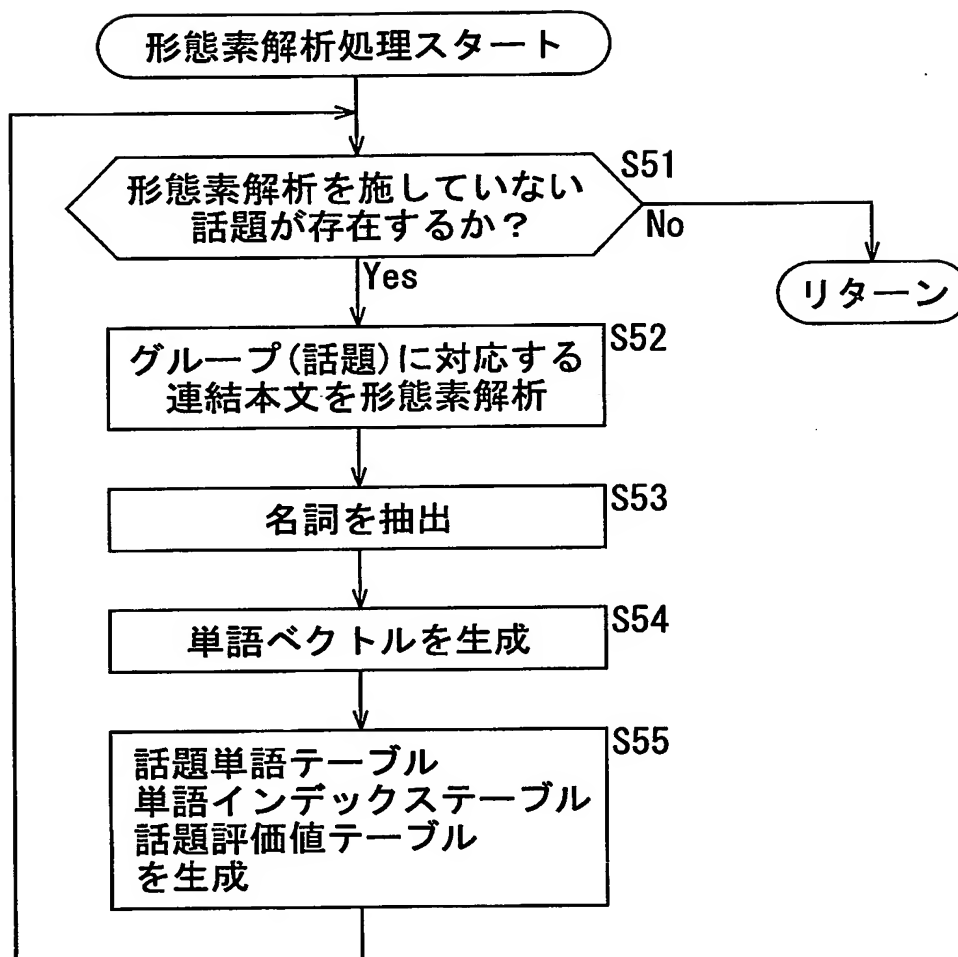
11/41

図12



12/41

図13



13/41

図14

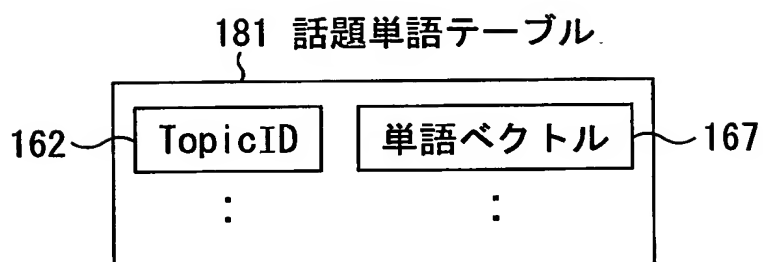


図15

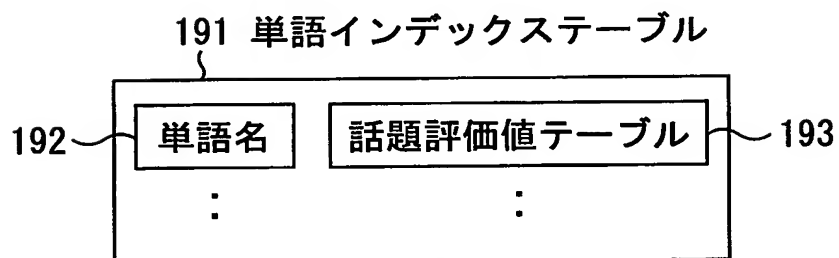


図16

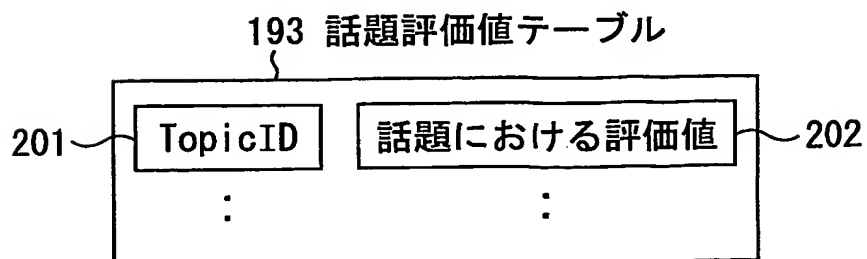
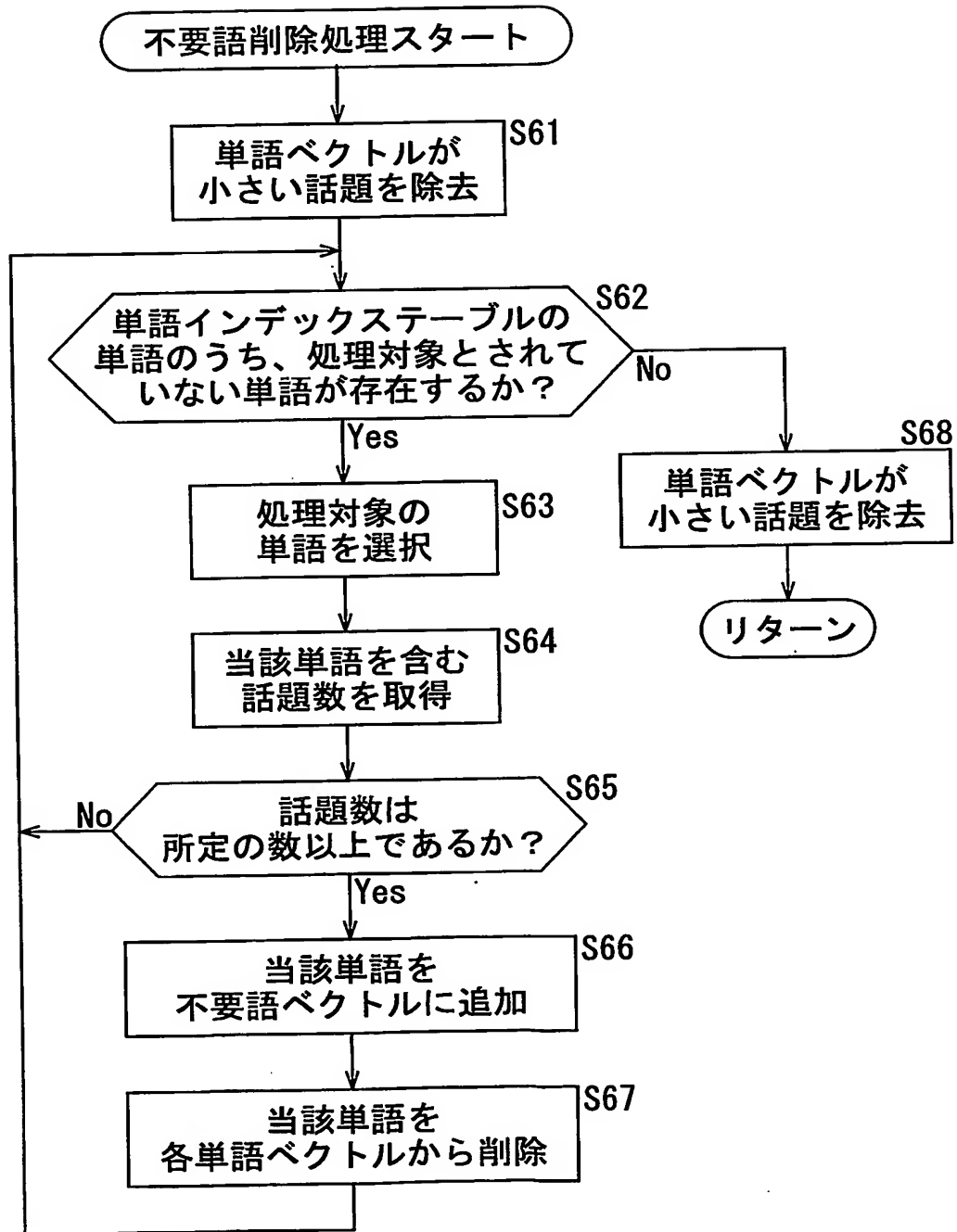
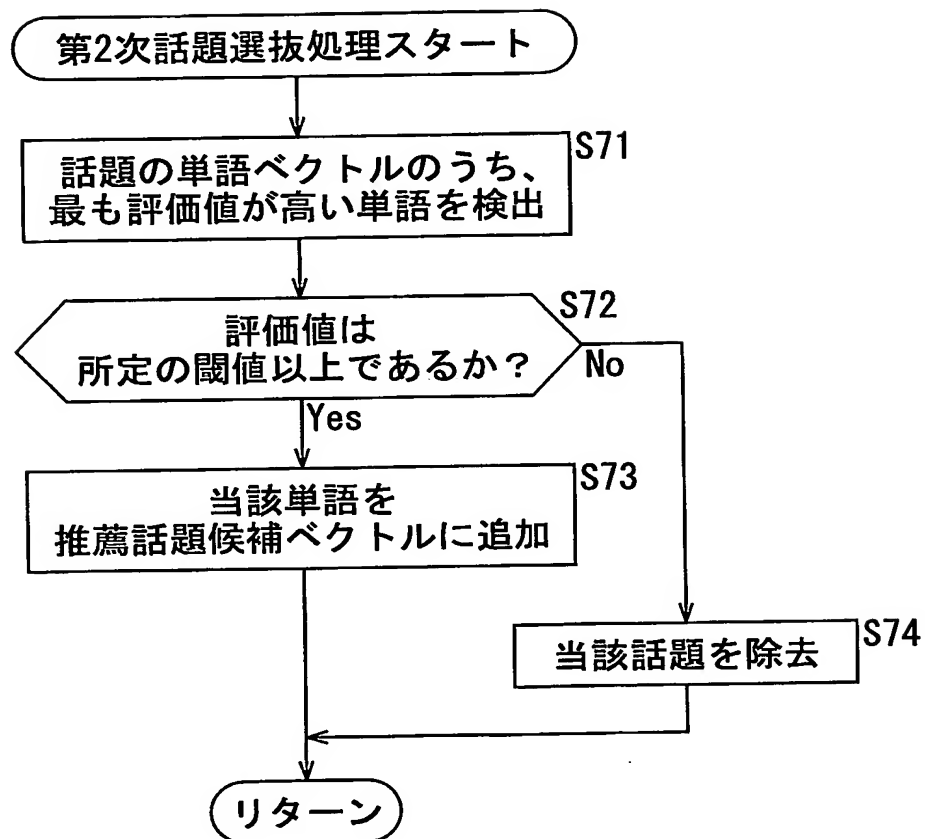


図17



15/41

図18



16/41

図19

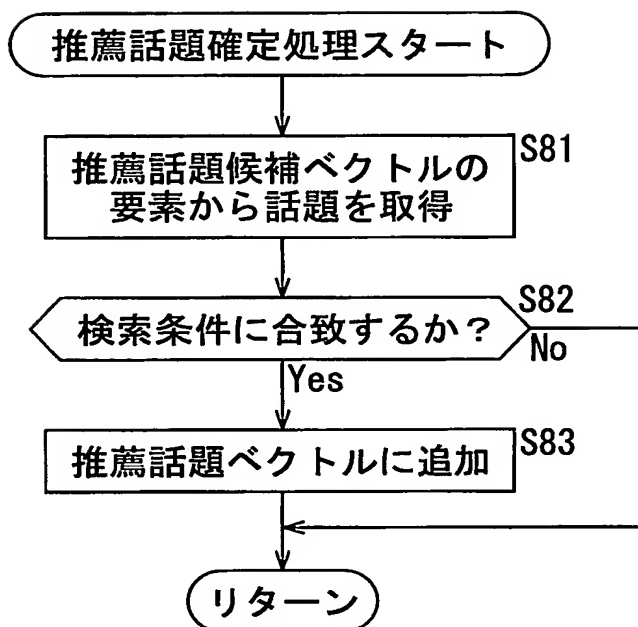
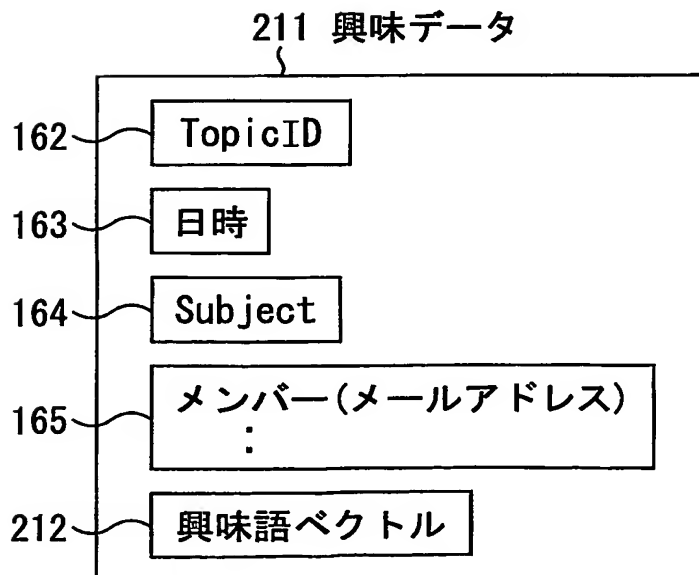


図20



17/41

図21

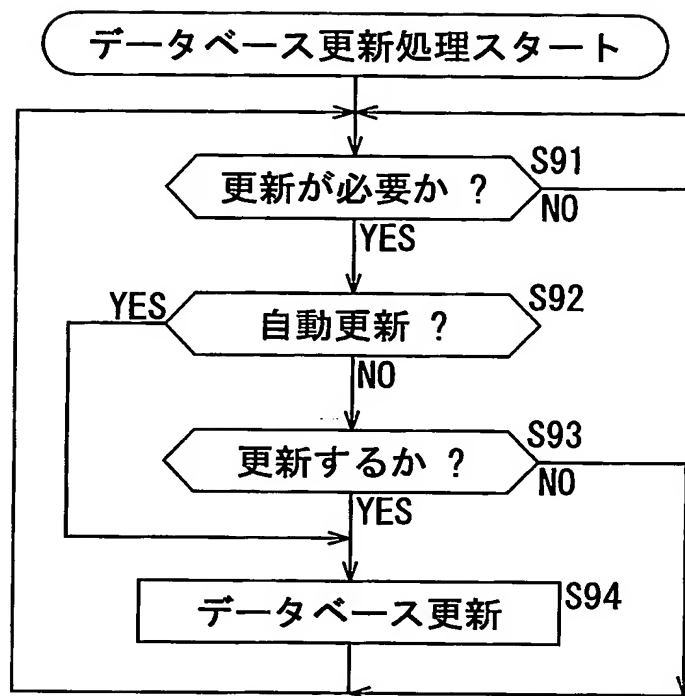
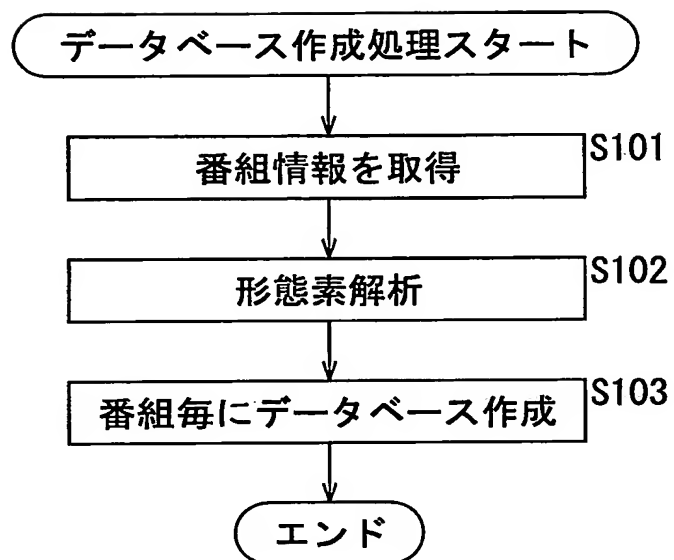


図22

- ☐ URLが ☐ ▼ 回行われたメールグループは更新必要と判断する
- ☐ 推薦URLを使い切ったメールグループは更新必要と判断する
- ☐ 最終更新から ☐ ▼ 日経過したメールグループは更新必要と判断する
- ☐ 新着メールの分析により、検索語が変更されたメールグループは更新必要と判断する

19/41

図23



20/41

図24

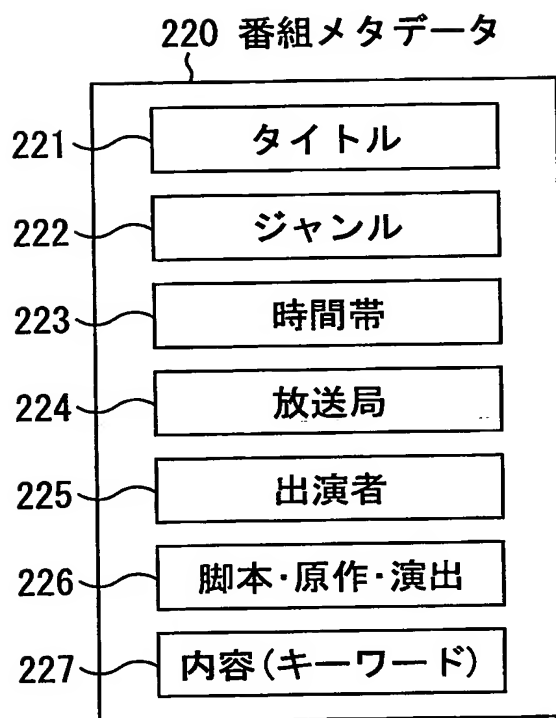
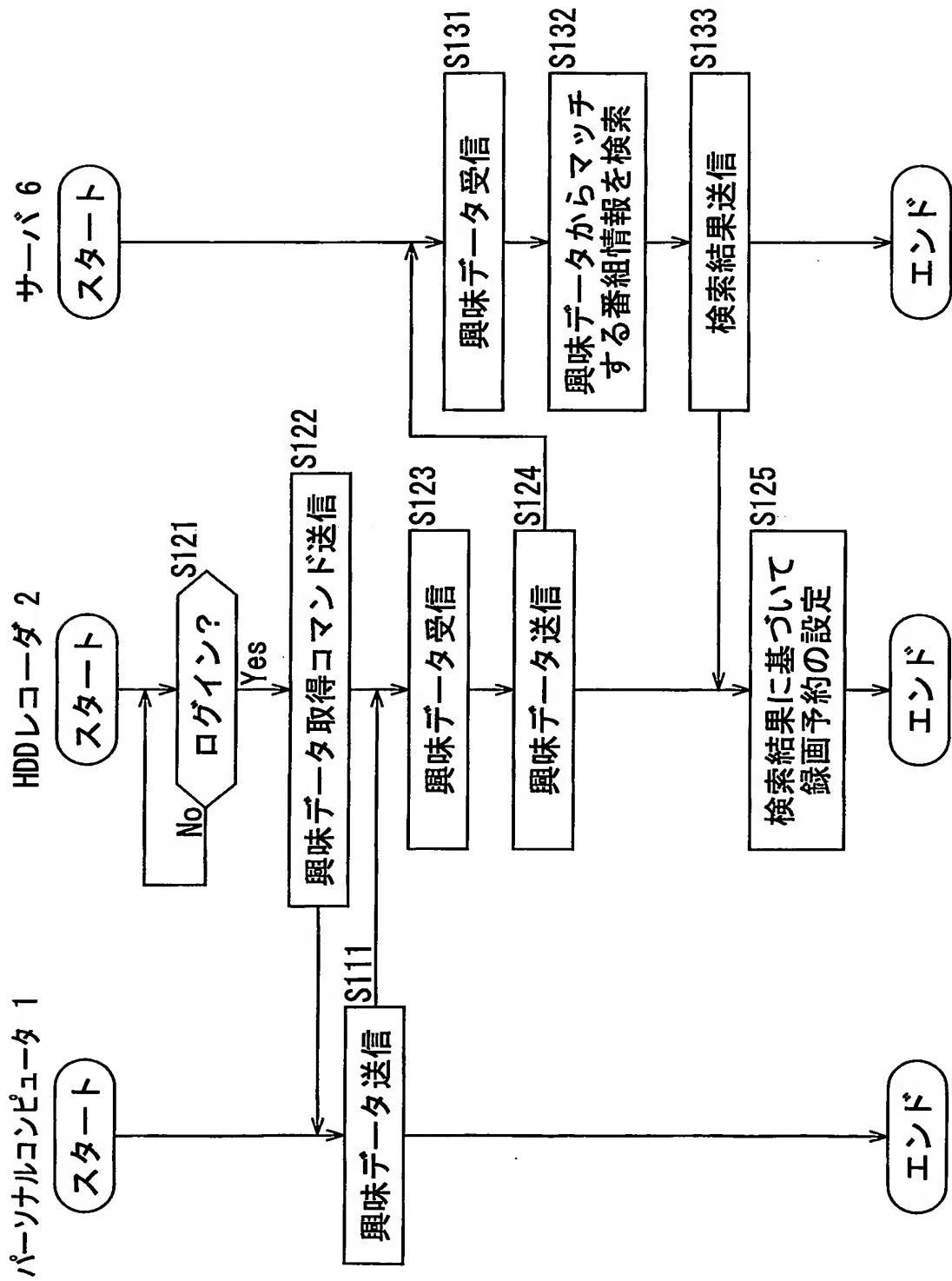
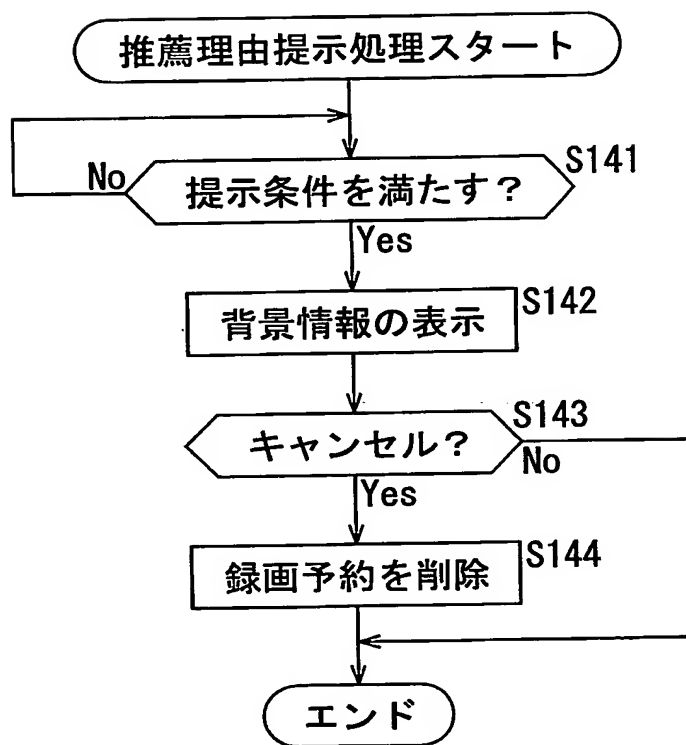


図25



22/41

図26



23/41

図27

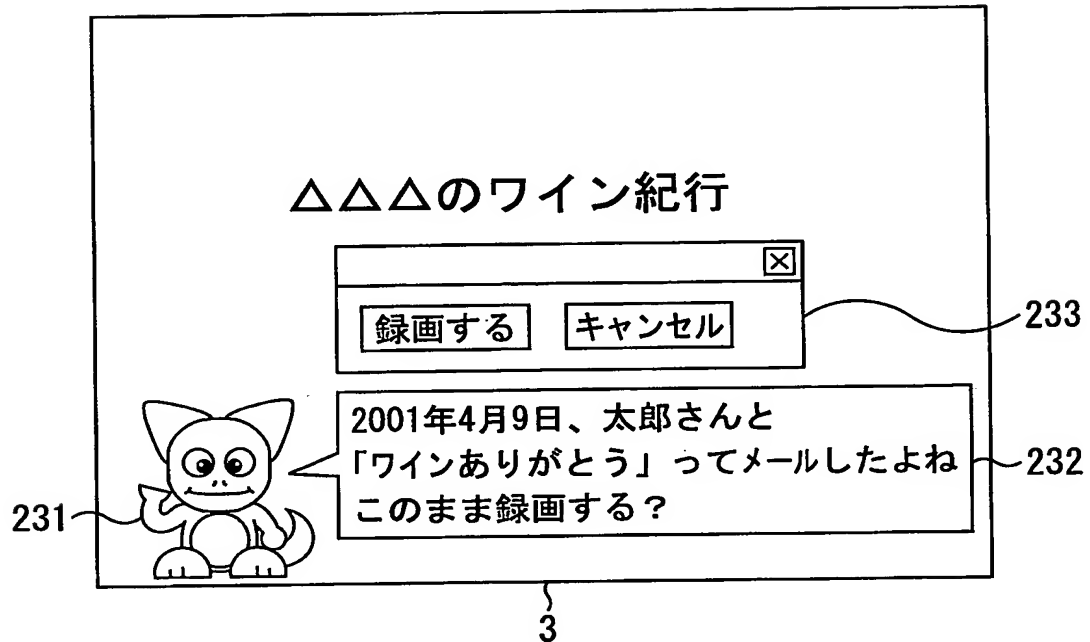


図28

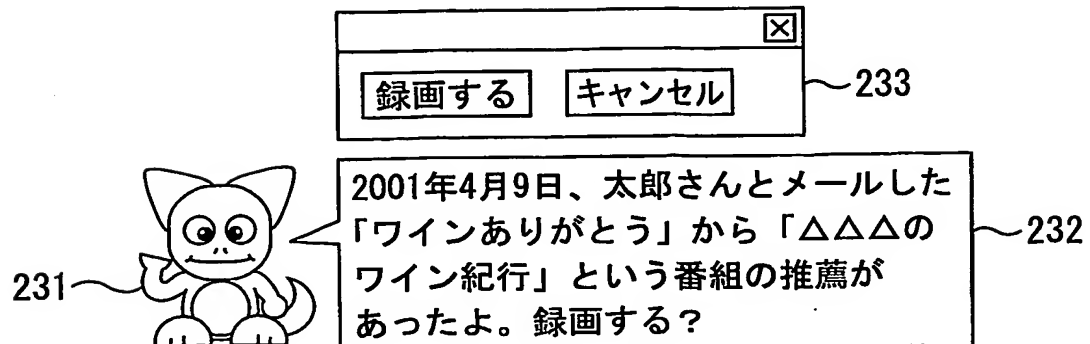
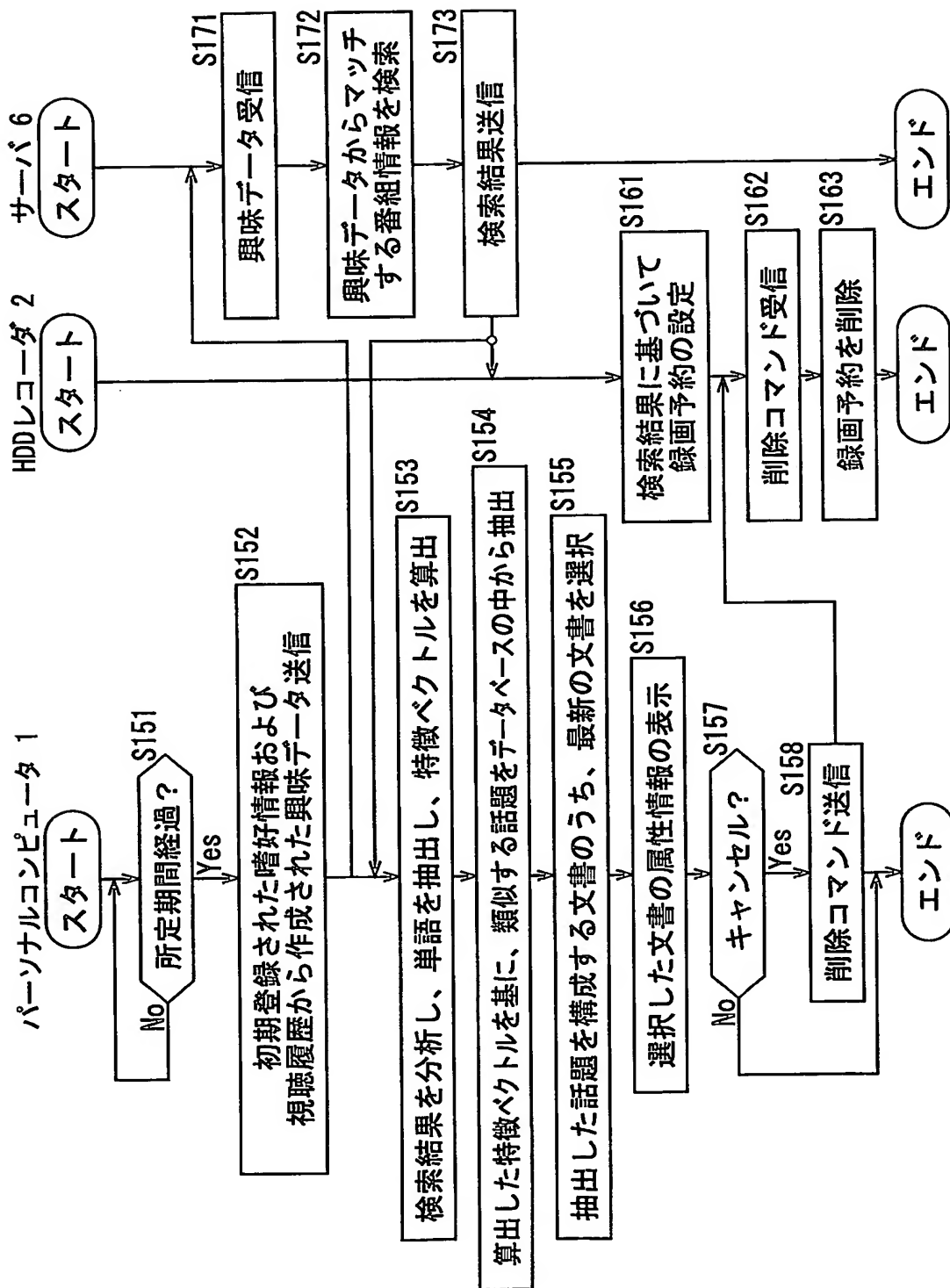
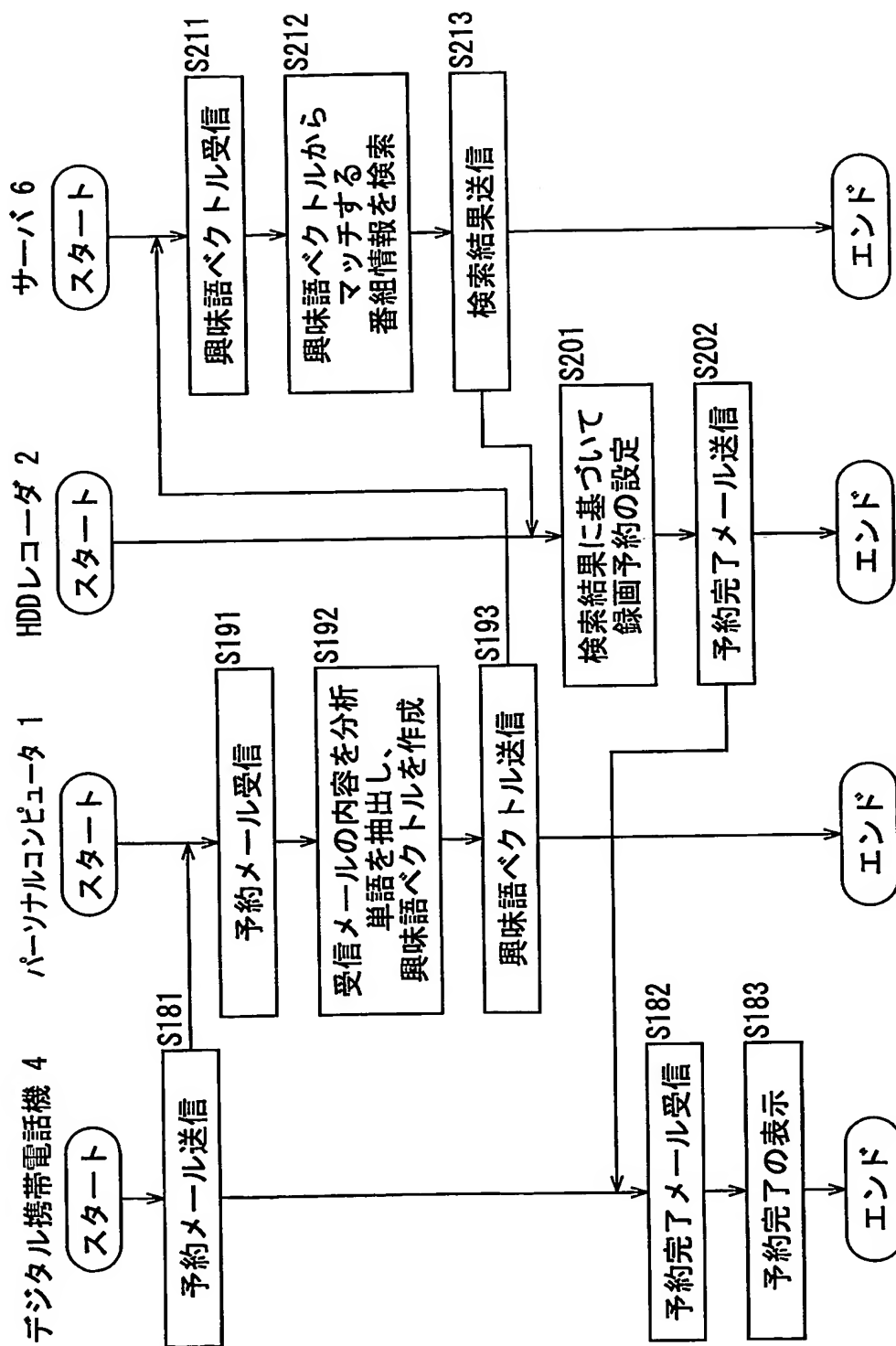


図29



25/41

図30



26/41

図31

To: * * @aa. co. jp
From: ○○○@bb. co. jp
Subject: 予約メール

今日、音楽番組を録画して。
特にクラシックものは絶対。
あればジャズ番組も。
あとワールドカップ情報も
お願い。

～241 予約メール

図32

To: ○○○@bb. co. jp
From: △△@cc. co. jp
Subject: 予約完了メール

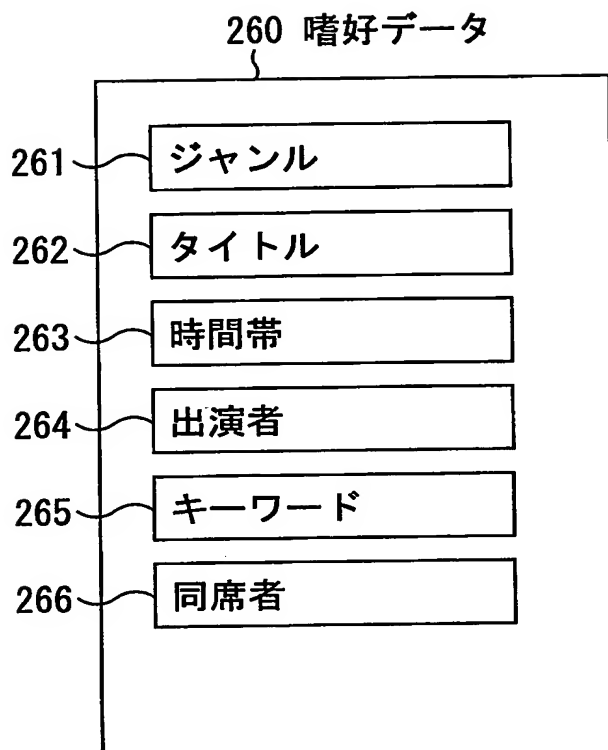
下記の番組を録画しました

- 1 4CH 19:00～20:00
「ワールドカップハイライト」
- 2 3CH 21:00～21:54
「×××クラシック」

～251 予約完了メール

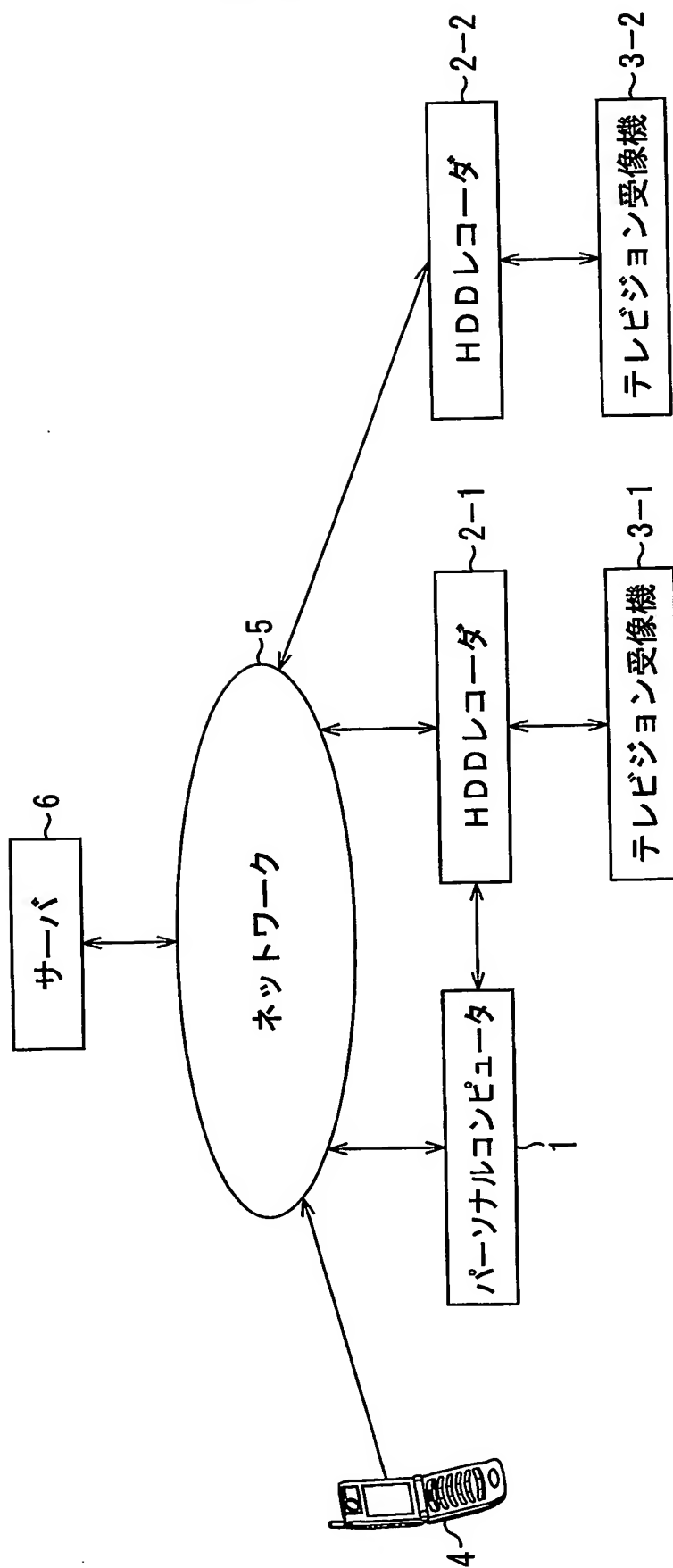
27/41

図33



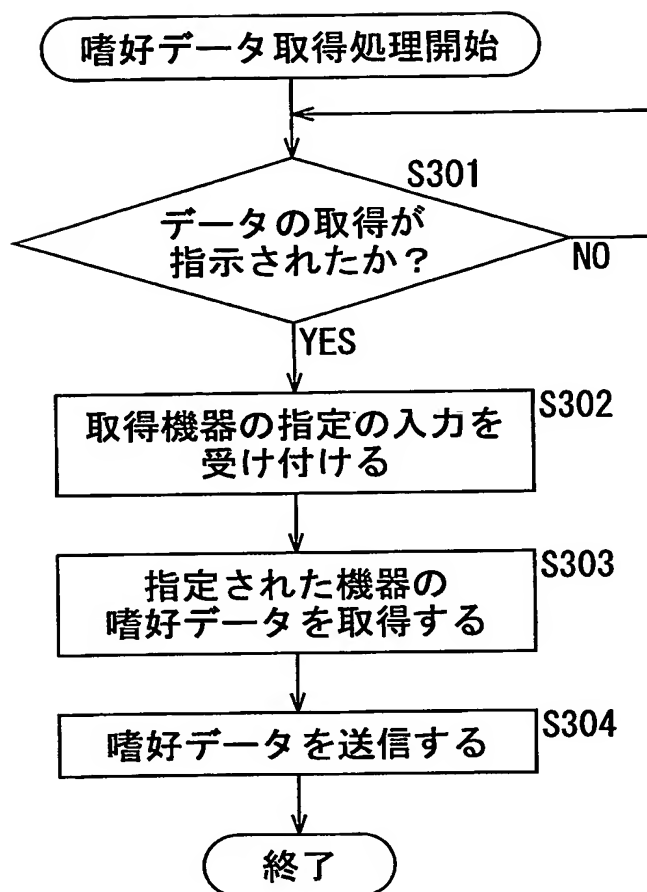
28/41

図34



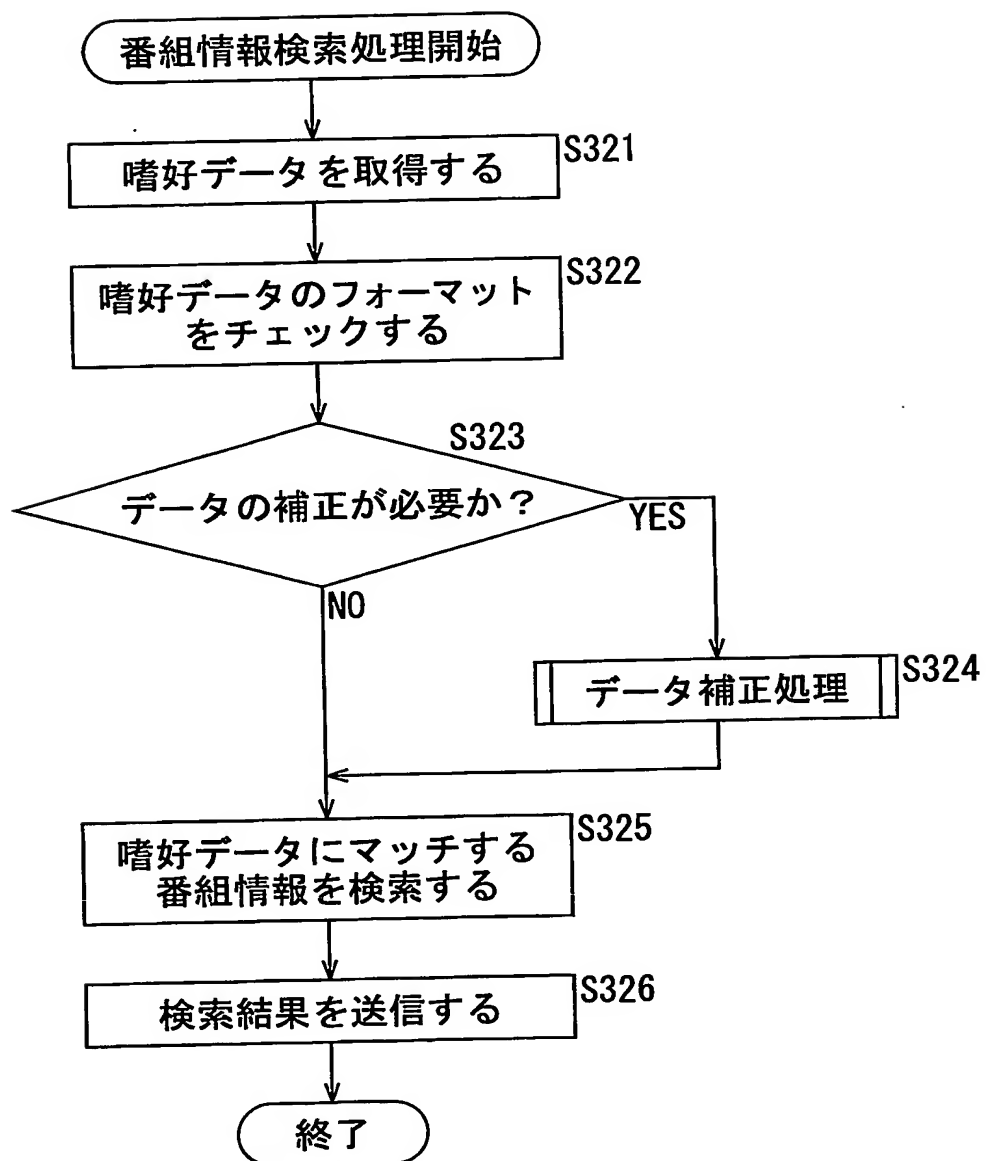
29/41

図35



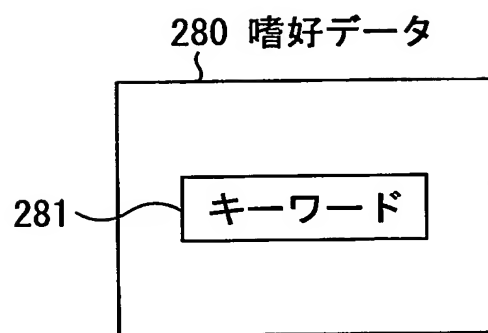
30/41

図36



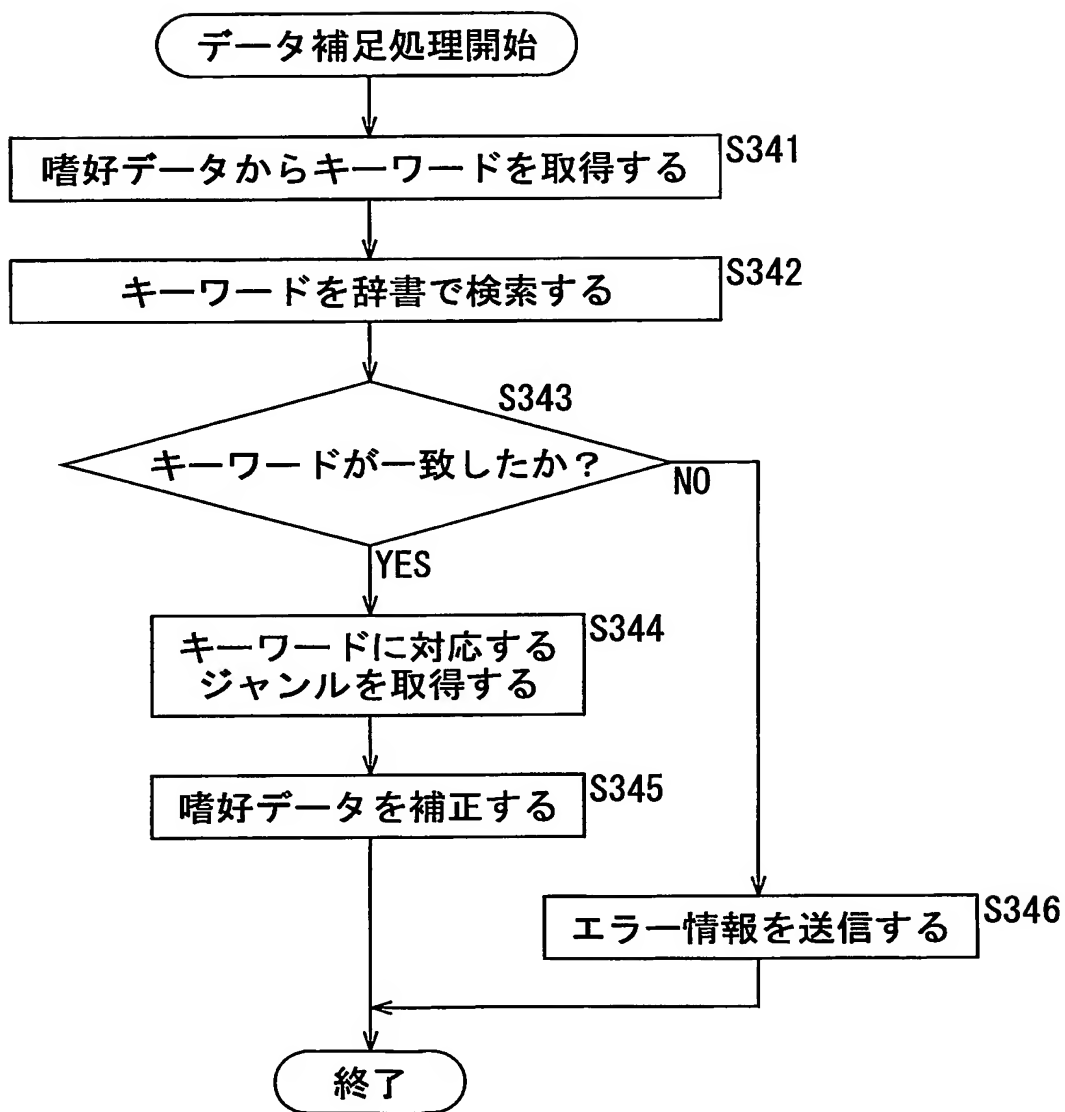
31/41

図37



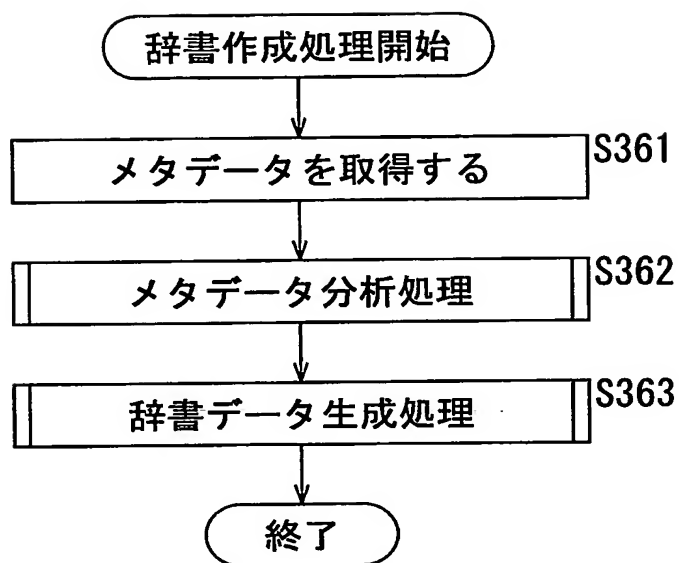
32/41

図38



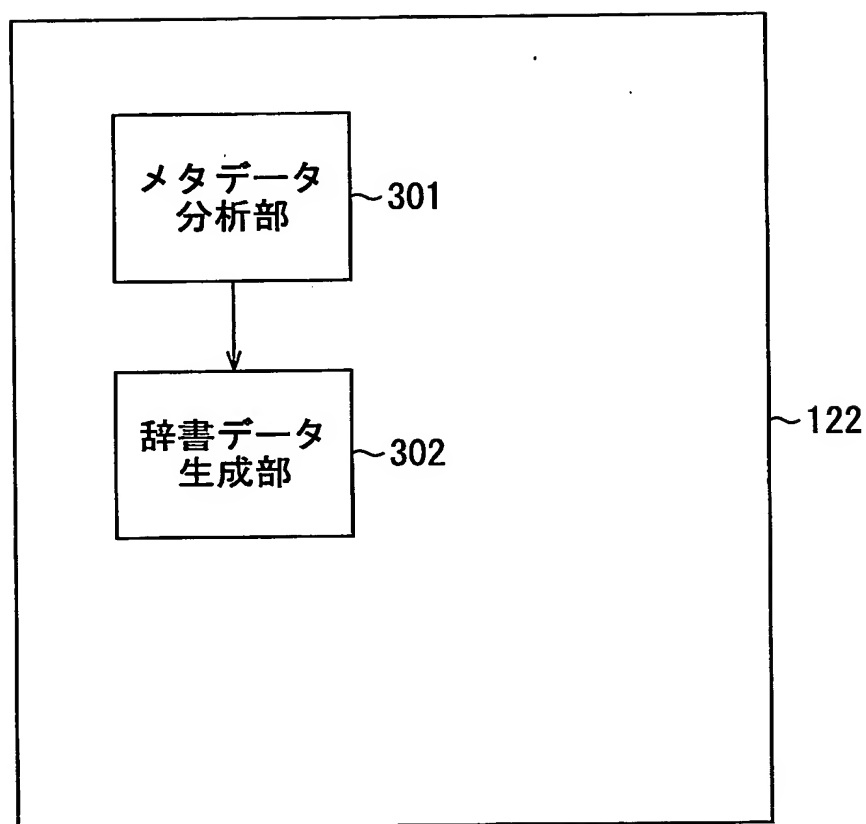
33/41

図39



34/41

図40



35/41

図41

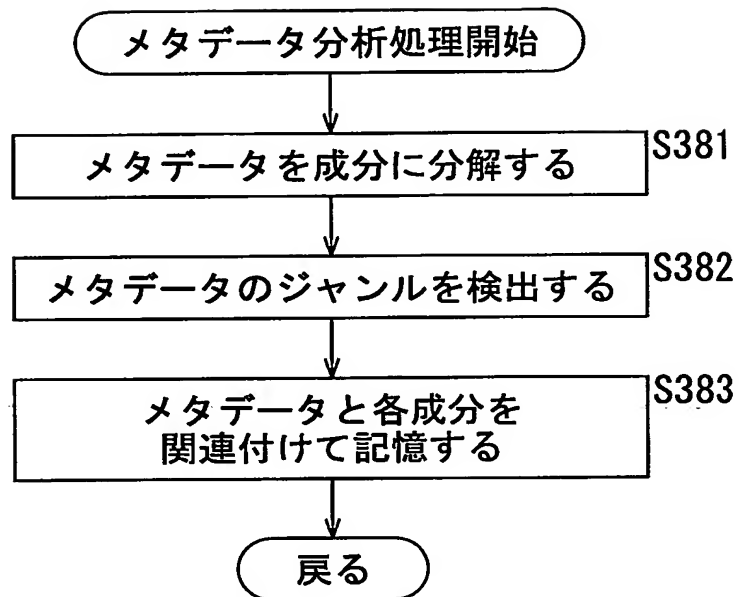


図42

成分	ジャンル	放送局	放送時間帯	出演者	キーワード	...
No1	料理	TAS	昼	AAA	レシピ、材料、作り方、文字放送、ステレオ	...
No2	生活情報	MHK	夜	BBB	レジャー、行楽地、子供、文字放送、ステレオ	...
No3	子供	MHK	朝	CCC	レジャー、子供、文字放送、ステレオ	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

図43

成分	ジャンル	放送局	放送時間帯	出演者	キーワード	...
No1-1	料理	TAS	昼	AAA	レシピ、材料、作り方、文字放送、ステレオ	...
No1-2	料理	MHK	昼	DDD	レシピ、材料、作り方、煮物、文字放送、ステレオ	...
No1-3	料理	MHK	朝	EEE	レシピ、材料、作り方、台所、文字放送、ステレオ	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

38/41

図44

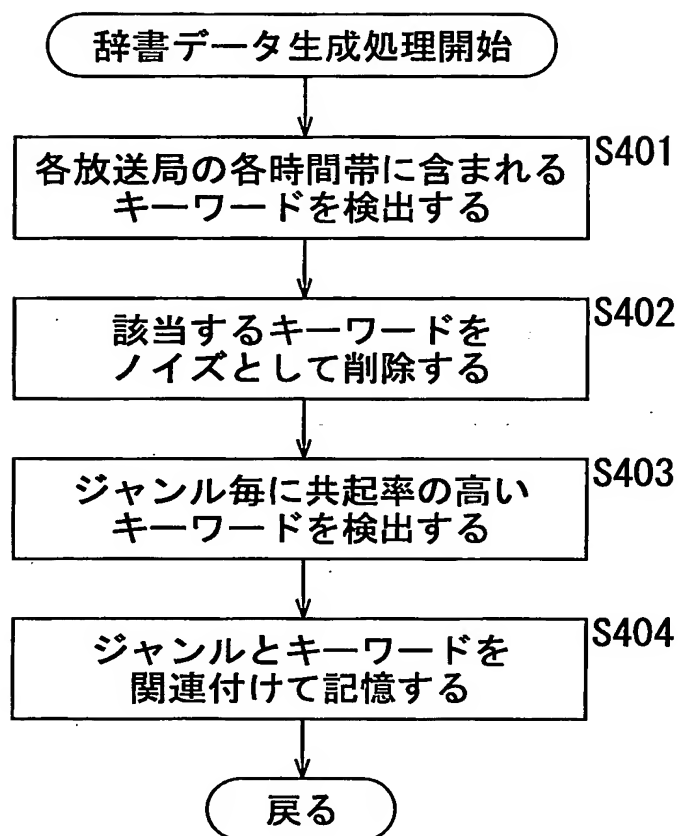
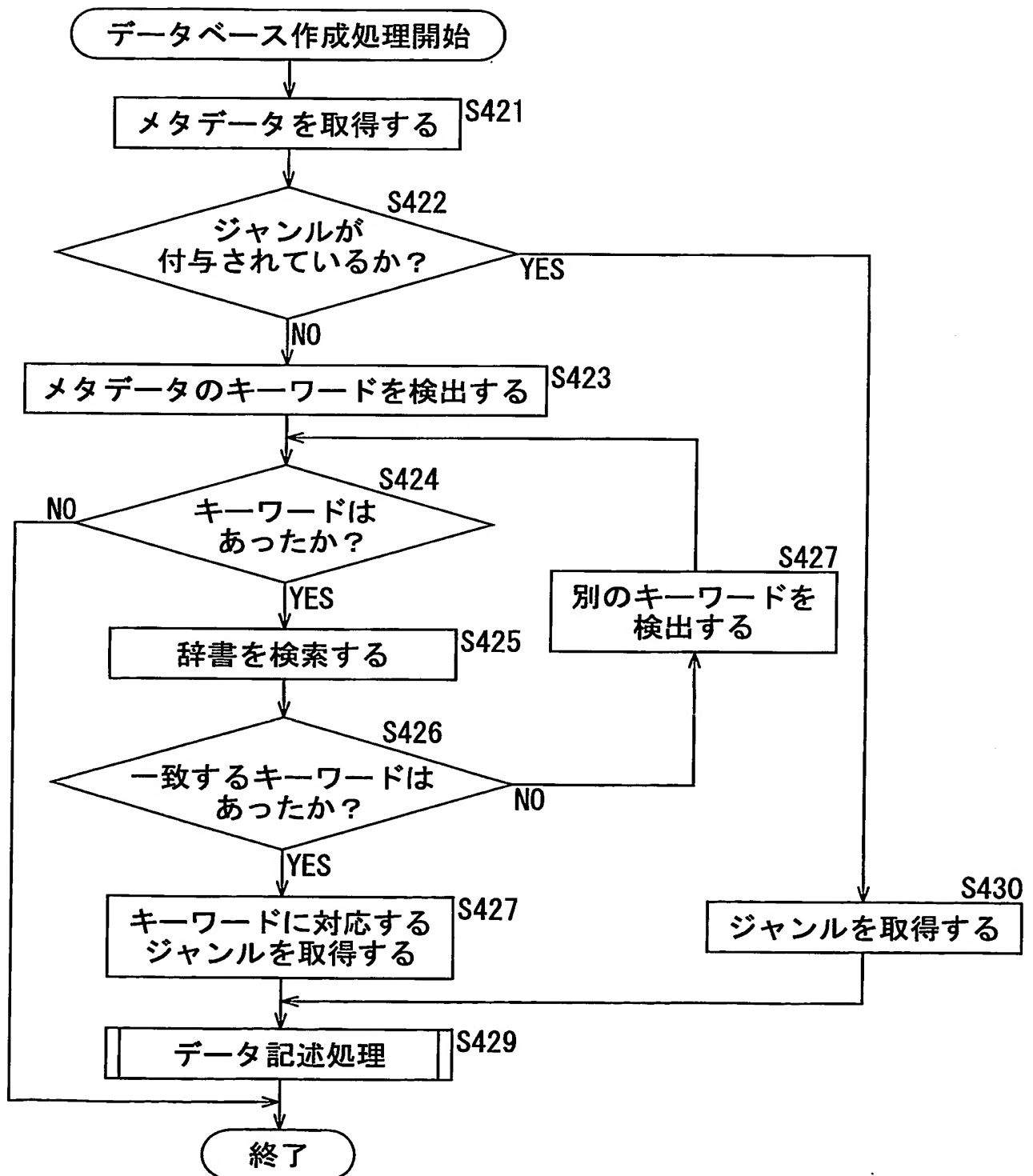


図45

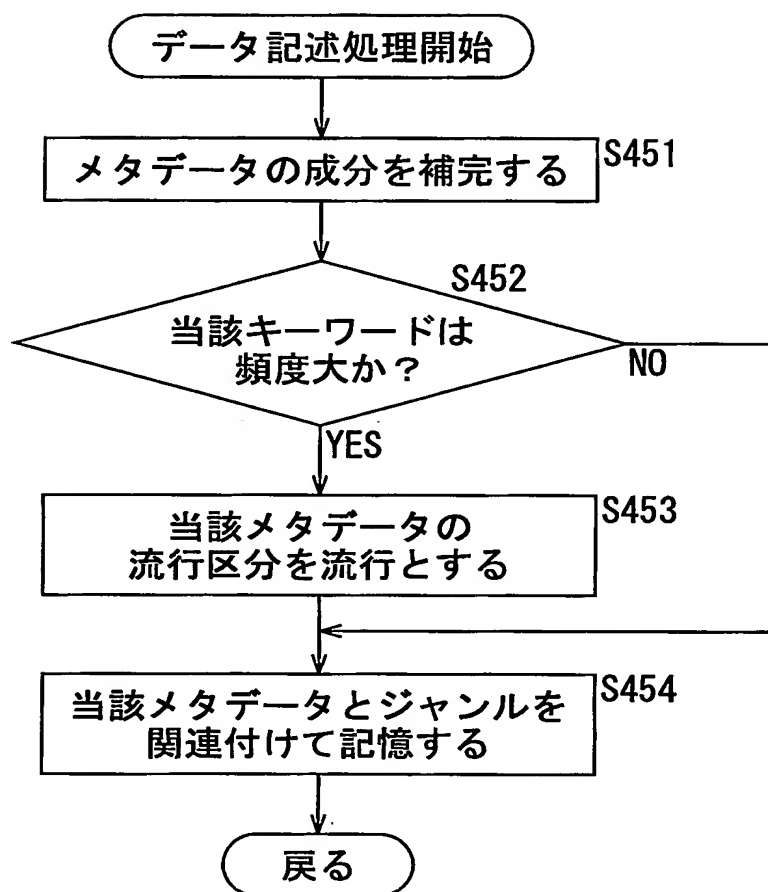
	キーワード	頻度／月	ジャンル	他の成分
1	レシピ	3	料理	...
2	材料	7	料理	...
3	作り方	2	料理	...
4	レジャー	15	生活情報、子供	...
5	行楽地	3	生活情報	...
6	子供	12	生活情報、子供	...
∴	∴	∴	∴	∴

図46



41/41

図47



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/03795

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl⁷ G06F17/30, H04N7/173

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl⁷ G06F17/30, H04N7/173

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2003
Kokai Jitsuyo Shinan Koho 1971-2003 Jitsuyo Shinan Toroku Koho 1996-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JICST FILE (JOIS), WPI, INSPEC (DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 10-257405 A (Mitsubishi Electric Corp.), 25 September, 1998 (25.09.98), Full text; Figs. 1 to 14 (Family: none)	1-54
Y	JP 2002-051287 A (Sony Corp.), 15 February, 2002 (15.02.02), Full text; Figs. 1 to 22 (Family: none)	1-54
Y	JP 11-007453 A (Mitsubishi Electric Corp.), 12 January, 1999 (12.01.99), Full text; Figs. 1 to 45 (Family: none)	1-54

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to
"A" document defining the general state of the art which is not	understand the principle or theory underlying the invention
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be
"L" document which may throw doubts on priority claim(s) or which is	considered novel or cannot be considered to involve an inventive
cited to establish the publication date of another citation or other	step when the document is taken alone
special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be
"O" document referring to an oral disclosure, use, exhibition or other	considered to involve an inventive step when the document is
means	combined with one or more other such documents, such
"P" document published prior to the international filing date but later	combination being obvious to a person skilled in the art
than the priority date claimed	"&" document member of the same patent family

Date of the actual completion of the international search
22 April, 2003 (22.04.03)

Date of mailing of the international search report
06 May, 2003 (06.05.03)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/03795

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2001-189896 A (Canon Inc.), 10 July, 2001 (10.07.01), Full text; Figs. 1 to 9F (Family: none)	1-54
Y	JP 2001-283101 A (Sony Corp.), 12 October, 2001 (12.10.01), Full text; Figs. 1 to 16 (Family: none)	2, 7, 12, 16, 24, 35, 42
Y	JP 2001-282830 A (Sony Corp.), 12 October, 2001 (12.10.01), Full text; Figs. 1 to 13 (Family: none)	2, 7-8, 16-17, 24, 35, 42, 50-54
Y	JP 2001-312515 A (Sony Corp.), 09 November, 2001 (09.11.01), Full text; Figs. 1 to 19 (Family: none)	2, 7-8, 16-17, 24, 35, 42, 50-54
A	JP 2001-275048 A (Victor Company Of Japan, Ltd.), 05 October, 2001 (05.10.01), Full txt; Figs. 1 to 5 (Family: none)	1-54
A	JP 2001-282831 A (Sony Corp.), 12 October, 2001 (12.10.01), Full text; Figs. 1 to 14 (Family: none)	1-54
A	JP 2001-057543 A (Sony Corp.), 27 February, 2001 (27.02.01), Full text; Figs. 1 to 15 (Family: none)	1-54
A	JP 2000-155764 A (Sony Corp.), 06 June, 2000 (06.06.00), Full text; Figs. 1 to 13 (Family: none)	1-54
A	JP 2000-341599 A (Matsushita Electric Industrial Co., Ltd.), 08 December, 2000 (08.12.00), Full text; Figs. 1 to 21 (Family: none)	1-54
A	JP 2002-077755 A (Sharp Corp.), 15 March, 2002 (15.03.02), Full text; Figs. 1 to 21 & EP 1189151 A2 & US 2002/0059180 A1	1-54

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/03795

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 07-135621 A (Matsushita Electric Industrial Co., Ltd.), 23 May, 1995 (23.05.95), Full text; Figs. 1 to 5 (Family: none)	1-54
A	JP 2001-312513 A (Sony Corp.), 09 November, 2001 (09.11.01), Full text; Figs. 1 to 22 (Family: none)	1-54

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G 06 F 17/30, H 04 N 7/173

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G 06 F 17/30, H 04 N 7/173

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2003年
日本国登録実用新案公報	1994-2003年
日本国実用新案登録公報	1996-2003年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICSTファイル (JOIS), WPI, INSPEC (DIALOG)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 10-257405 A (三菱電機株式会社) 1998. 09. 25, 全文, 第1~14図 (ファミリーなし)	1-54
Y	JP 2002-051287 A (ソニー株式会社) 2002. 02. 15, 全文, 第1~22図 (ファミリーなし)	1-54
Y	JP 11-007453 A (三菱電機株式会社) 1999. 01. 12, 全文, 第1~45図 (ファミリーなし)	1-54

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」 口頭による開示、使用、展示等に言及する文献
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」 同一パテントファミリー文献

国際調査を完了した日

22. 04. 03

国際調査報告の発送日

06.05.03

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

田川 泰宏

5M

4236

電話番号 03-3581-1101 内線 3597

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 2001-189896 A(キャノン株式会社) 2001. 07. 10, 全文, 第1~9F図 (ファミリーなし)	1-54
Y	JP 2001-283101 A(ソニー株式会社) 2001. 10. 12, 全文, 第1~16図 (ファミリーなし)	2, 7, 12, 16, 2 4, 35, 42
Y	JP 2001-282830 A(ソニー株式会社) 2001. 10. 12, 全文, 第1~13図 (ファミリーなし)	2, 7-8, 16-17, 24, 35, 42, 50- 54
Y	JP 2001-312515 A(ソニー株式会社) 2001. 11. 09, 全文, 第1~19図 (ファミリーなし)	2, 7-8, 16-17, 24, 35, 42, 50- 54
A	JP 2001-275048 A(日本ビクター株式会社) 2001. 10. 05, 全文, 第1~ 5図 (ファミリーなし)	1-54
A	JP 2001-282831 A(ソニー株式会社) 2001. 10. 12, 全文, 第1~14図 (ファミリーなし)	1-54
A	JP 2001-057543 A(ソニー株式会社) 2001. 02. 27, 全文, 第1~15図 (ファミリーなし)	1-54
A	JP 2000-155764 A(ソニー株式会社) 2000. 06. 06, 全文, 第1~13図 (ファミリーなし)	1-54
A	JP 2000-341599 A(松下電器産業株式会社) 2000. 12. 08, 全文, 第1~ 21図 (ファミリーなし)	1-54
A	JP 2002-077755 A(シャープ株式会社) 2002. 03. 15, 全文, 第1~21図 &EP 1189151 A2 &US 2002/0059180 A1	1-54
A	JP 07-135621 A(松下電器産業株式会社) 1995. 05. 23, 全文, 第1~5 図 (ファミリーなし)	1-54
A	JP 2001-312513 A(ソニー株式会社) 2001. 11. 09, 全文, 第1~22図 (ファミリーなし)	1-54